

nature

THE INTERNATIONAL WEEKLY JOURNAL OF SCIENCE



The epigenetic
trait that can
ruin an oil
palm crop

PAGES 466 & 533

Weed out *Bad Karma*

PRIVACY

OLD TAX FILES WANTED

Making administrative data
accessible and anonymous

PAGE 440

ENERGY

DUAL-USE FUEL CELLS

A smart way to boost
electrical power grids

PAGE 447



CAREERS

USE YOUR IMAGINATION

Sci-fact as an
entrée into sci-fi

PAGE 553

NATURE.COM/NATURE

24 September 2015 £10

Vol. 525, No. 7570



39p

THIS WEEK

EDITORIALS

STEM CELLS Curtain comes down on scientific controversy **p.426**

WORLD VIEW End the unfair racket of academic jobs for the boys **p.427**

DECEPTION Orchid shape and smell fools amorous wasps **p.429**



In the name of beauty

The ugly truth is that the plastic microbeads found in many skin scrubs and other personal-care products are a serious pollutant of the marine environment. They should be phased out rapidly.

A beautiful woman comes into focus. What makes her skin glow so? Why, she says, she uses Aveeno's Positively Radiant skin-brightening daily scrub for "naturally beautiful results".

What is not clear from this advertisement is that the "gentle exfoliators" in the product promoted by Jennifer Aniston are minuscule beads of plastic. When Aniston, or those she inspires to follow her, rinse the scrub down the drain, many of the beads end up in the sea, where they will persist indefinitely. This is unnecessary, damaging and must stop.

Others agree, and the face scrub, along with hundreds of other products, including toothpastes, may not be long for this world. On 10 September, the California Legislature sent a bill (AB 888) to the state's governor, Jerry Brown, that would ban the inclusion of spheres of polyethylene, polypropylene and other plastics less than 5 millimetres across in personal-care products after 2020.

If signed into law, the bill will prevent trillions of plastic beads from being rinsed down the drain. Not all of these make it to the sea — wastewater treatment plants can sift out 90% of them — but the problems caused by the remaining millions are considerable. (Meanwhile, beads trapped in 'sludge' at the plants do not disappear. Plenty are sprayed on crops, from where they escape to rivers and lakes.)

In a paper published on 3 September, aquatic-health researcher Chelsea Rochman at the University of California, Davis, and her colleagues estimate that 8 trillion microbeads per day are emitted into aquatic habitats in the United States alone (C. M. Rochman *et al. Environ. Sci. Technol.* <http://doi.org/7sw>; 2015).

The beads are more pernicious than mere litter. Roughly the size of many plankton species, they are eaten by marine creatures. One study in 2014 saw them consumed by several taxa of zooplankton, including mysid shrimps, copepods, rotifers and ciliates (O. Setälä *et al. Environ. Pollut.* **185**, 77–83; 2014). Some of these are then eaten by larger creatures, and toxic chemicals in the plastics, as well as other toxic chemicals that adhere to plastic particles, accumulate in fish — which might end up on our dinner tables.

California would not be the first place to pass a microbead ban, but as the world's seventh- or eighth-largest economy, its move would carry weight. Just as in automotive-fuel efficiency standards or flammability requirements on furniture, where California goes, other places in the United States and elsewhere follow. The California bill is also stronger than many before it. It does not include a common loophole allowing for the use of 'biodegradable' beads — which are unlikely to truly degrade anywhere except in an industrial composter.

California legislators have made the right call, but the phase-out period is too long. No luminous complexion is worth the wholesale pollution of Earth's oceans. Consumer-goods giant Unilever says that it has already removed microbeads from all of its scrubs and washes. And there are plenty of well-tested alternative exfoliants, including nut shells, sand and sugar. So why wait five years to stop polluting?

While bans and phase-outs slowly take effect, the Beat the Microbead campaign, funded by Dutch non-governmental organizations the Plastic Soup Foundation and the North Sea Foundation, has created an app for consumers who want to avoid contributing to the problem. A few clicks can confirm whether the tempting scrub in the pharmacy aisle contains the beads. This is helpful in the short term, but ultimately the onus of responsibility should not be on the consumer.

"No luminous complexion is worth the wholesale pollution of Earth's oceans."

Microbeads are not the only source of microplastic in the oceans. Tiny plastic pellets used in making plastic items spill into the sea; plastic bags and bottles break down over time. On almost any beach on Earth, the sand carries tiny, bright grains of plastic.

And macroplastics remain a serious problem. A study published last month estimated that around 90% of seabirds have plastic in their bellies (C. Wilcox *et al. Proc. Natl Acad. Sci. USA* <http://doi.org/7dv>; 2015). Some birds mistake shopping bags for jellyfish; others confuse cigarette lighters and pen caps with prey and fly home to feed them to their chicks.

The consequences of this ubiquitous plastic for marine species, marine ecosystems and human health remain areas of active research. But the public and policymakers need not wait for detailed results before taking action. Banning microbeads will not solve the plastic-pollution problem, but it is an easy start. Jennifer Aniston and the millions of other people who wash their faces with plastic can still look radiant without feeding their skincare regime to copepods. The alternative is to forever blush with shame. ■

Power play

The replacement of mitochondria does not signal ethical problems.

Galloping advances in genetic and stem-cell technologies raise the inviting prospect that some devastating diseases could be treated. Conditions caused by natural mutations might be avoided by judicious genome tinkering to set things right for the next generations. But 'inviting' does not always mean 'advisable'.

The United Kingdom last week released new draft guidelines for one such treatment — mitochondrial replacement (see go.nature.com/thcouy). The guidelines are scheduled to come into force next month, when clinics in Britain will be allowed to offer the treatment. Not everybody agrees that this inviting idea is advisable. As such, it is timely to consider the ethical and technical matters at stake.

At issue is not the nuclear genome, which contains the blueprint of an entire organism, but the genomes in our mitochondria — the small, energy-generating organelles in most of our cells. The often-overlooked mitochondrial genome contains only a few dozen genes, but it deserves as much respect as its much larger room-mate, which contains some 20,000. The impacts of an unfortunate mitochondrial mutation range from an inability to exercise hard to very serious, albeit rare, diseases.

Mitochondrial replacement involves replacing diseased mitochondria with fresh, healthy ones. This requires involving a third person beyond the parents — a woman to donate an egg to the process that contains only healthy mitochondria (hence ‘three-person embryo’).

The procedure does not alter the mitochondrial genome. But on the basis of animal experiments, some biologists claim that foreign mitochondrial genes might interfere with the expression of the nuclear genome in unpredictable, and perhaps dangerous, ways (see page 444).

These concerns were brought up during the consultation process with scientists and the public carried out by the UK Human Fertilisation and Embryology Authority (HFEA) before the UK Parliament voted in February in favour of allowing the procedure. Far from being rushed, as some claim, the consultation was done over many years and was judged as a fair public-engagement exercise by independent experts who monitored the process.

The HFEA believes that the problems seen in organisms such as flies and mice would not be repeated in humans — in the main because they have not shown up in children of mixed-race couples in which the mitochondrial DNA of the mother and the nuclear DNA of the father are likely to be the most distant. This point helps to address ethicists’ worries that unanticipated problems in children born following mitochondrial

replacement could be passed on through the generations.

Other ethical concerns about the UK move can be summarized as anxiety over a possible slippery slope to full-scale germline manipulation to address a broader range of conditions. These concerns are heightened by advances in gene-editing techniques such as CRISPR/Cas9.

Last week’s release of the HFEA regulations should dispel fears of a slippery slope. Applications are narrow and oversight is strict. The

“The HFEA regulations should dispel fears of a slippery slope.”

agency decided to allow mitochondrial replacement only to avoid serious diseases, and not for the attempted treatment of infertility. (Some clinics in Canada have offered the procedure in the belief that a shot of fresh, young mitochondria may somehow invigorate eggs from older women, but there is little scientific evidence for this.)

The regulations explicitly exclude the editing of the nuclear or mitochondrial genome. Licences will be given only to centres whose competence has been approved, and even then, these centres will have to seek separate approval for each patient. Licensed centres will be obliged to put a process in place to monitor the clinical follow-up of children born following mitochondrial donation, providing that parents agree.

Scientists estimate that the number of women likely to be eligible for the procedure will be around 150 per year in Britain and about 800 in the United States, where the Institute of Medicine is carrying out a similar consultation for the US Food and Drug Agency, which will be responsible for licensing it. The United Kingdom has made an advisable step forward that serves as a useful invitation for all to follow. ■

STAP revisited

Reanalysis of the controversy provides a strong example of the self-correcting nature of science.

This week, *Nature* revisits one of the most controversial scientific episodes in recent years: the now-retracted discovery of a claimed new way to reprogram cells, stimulus-triggered acquisition of pluripotency (STAP). On our website we publish two Brief Communications Arising (BCAs) that relate to the retraction. And on page 469 we publish a related Review on pluripotency.

One BCA details the efforts made by many laboratories to reproduce the STAP phenomenon without success (A. De Los Angeles *et al.* *Nature* <http://dx.doi.org/10.1038/nature15513>; 2015). The other presents the results of a genomic analysis of the claimed STAP cells, performed as part of a 2014 investigation by Japan’s RIKEN institute but not previously published (D. Konno *et al.* *Nature* <http://dx.doi.org/10.1038/nature15366>; 2015). Using sequencing-based approaches, this analysis shows that all of the claimed STAP cell lines were contaminated with embryonic stem cells, and that this contamination affected the results. De Los Angeles and colleagues’ BCA also includes an analysis of sequencing results published in the original papers, and reaches similar conclusions regarding contamination.

The Review, written by a collaboration of leading scientists who work with pluripotent stem cells, offers a state-of-the-art summary of the field, and provides a checklist that researchers can use to determine whether a cell has pluripotent capacity.

Why is *Nature* publishing these pieces? The main reason is to update the scientific record. The wording of the STAP retraction notices left open the possibility that the phenomenon was genuine. It said: “Multiple errors impair the credibility of the study as a whole and we are unable to say without doubt whether the STAP-SC phenomenon is real.” The two BCAs clearly establish that it is not.

It is also important to recognize and highlight the community-driven effort to reproduce the findings. The negative results of some of these efforts were made public informally during the controversy, but for some lengthy experiments this was not possible. Science-in-the-making can be made public immediately. But, ultimately, reproducibility efforts should be peer reviewed.

Another reason why *Nature* has chosen to publish this trio of pieces is to address some of the indirect questions posed by the high-profile controversy, which provoked discussions in both the stem-cell field and the broader research community. The Review, in particular, is intended to offer guidance from the community to help researchers, editors and reviewers to decide how best to evaluate future claims as well as how to view those already published in the scientific literature. Comparing the genotypes of reprogrammed pluripotent stem cells with those of parental cells, it points out, can check their provenance.

The stem-cell field holds enormous promise for therapy. As a result, all claims of considerable importance should be verified with utmost care before being made public. The Review suggests that such claims in the field of reprogramming and pluripotency should be demonstrated in more than one experimental model, and encourages their independent replication.

Nature will endeavour to help the field to achieve its promise, and is looking at ways to support and encourage this reproducibility enterprise. For example, we ask authors to include more details about the methods developed in their studies. We strongly encourage our authors to deposit step-by-step protocols on freely accessible platforms, such as Protocol Exchange (www.nature.com/protocolexchange) — this may be requested for extraordinary claims, at the editor’s discretion. We encourage our authors to verify the origin of the cell lines they use, as we do for cancer cell lines (see *Nature* 520, 264; 2015).

➤ NATURE.COM
To comment online,
click on Editorials at:
go.nature.com/xhunqv

The Review concludes: “Science is ultimately a self-correcting process where the scientific community plays a crucial and collective role.” In this case, the stem-cell community has excelled in that role and should be congratulated. ■



Make academic job advertisements fair to all

Too many university posts are given to men without proper competition, says Mathias Wullum Nielsen.

It is well known that women are under-represented in senior science and research positions. This is true even in Denmark, which has long been considered one of the most advanced societies when it comes to gender equality. Although stories of sexism in science often focus on explicit bias, more-subtle factors are widely influential too.

Universities like to think of themselves as meritocracies. Indeed, one of the arguments used against programmes that aim to proactively promote the careers of women scientists is that scientists must be recruited on talent alone. When criticized over the appointment of (another) male scientist to a senior role, universities often respond by pointing to rules and policies about how vacancies invite all to apply.

I carried out an analysis that raises some troubling questions about how closely universities follow these principles (M. W. Nielsen *Sci. Public Policy* <http://doi.org/7q6>; 2015). In the decade to 2013, about one-fifth of associate- and full-professor positions at Aarhus University, one of the largest in Denmark, were filled through a 'closed' recruitment procedure: no advertisement and usually just a single applicant. The share of female candidates for such positions is particularly low — just 12% of applicants for full professorships were women.

With 'open' recruitment, the proportion of female applicants for full-professor roles rises to 23%. But a significant proportion still attracted only a single applicant, suggesting that the adverts were being written to target a specific candidate, or timed to fit their career progression. Evidence suggests that similar practices, to various extents, are common at other Danish research institutions and abroad.

Despite institutional efforts to make recruitment more robust and transparent, and a 2008 Danish ministerial decree that "professorships must be advertised internationally, except under special circumstances", my analysis shows that the use of closed recruitment at Aarhus University increased from 8% of tenured appointments in 2004–08 to 30% in 2009–13.

Such appointments do not usually break any rules; loopholes can be exploited in most cases. But the numbers suggest a lack of institutional commitment to overarching organizational and legal stipulations. And they confirm what most academics may already suspect: rising in the ranks is not a question merely of what you know, but of who you know.

This puts women at a particular disadvantage. Academic advancement through back-door hiring largely depends on reputation and visibility to the local gatekeepers, and women lose out under such procedures for two reasons.

First, women have been shown to have

weaker personal ties to the core of the concentric circles of academic networks, making them less visible to decision-makers. Second, scholars have argued that male decision-makers' desire for organizational certainty and their attraction to candidates with whom they share values and behaviour, create subtle and often unconscious practices of 'male closeness' and 'gender homophily' (preference for someone similar to oneself).

Gender scholars have previously shown that discrimination may be particularly prevalent in organizations that pride themselves on being meritocratic. Strong institutionalized beliefs in meritocracy are more likely to discourage people from paying attention to their own implicit biases and prejudices.

Sure enough, department heads seem unaware of the implications.

I interviewed 24 at Aarhus about whether and how issues of gender influence their recruitment and selection practices. I was told frequently that "gender doesn't play any part", "for us it's all about getting the best" and "what we look at is quality".

As Liisa Husu, a gender-studies researcher at Örebro University in Sweden, has pointed out, the myriad disadvantages facing female academics often operate as "non-events" (*Nature* **495**, 35–38; 2013). Women are not being taken into account, encouraged or asked along to the same extent as their male colleagues. Seen as separate occurrences, such non-events may seem harmless. But just as academic success will often accrue to the already successful, marginal disadvantages accumulate over time through self-reinforcing effects, with clear implications for gender stratification.

With an interest in addressing the gender-equality challenge, Aarhus University provided the data for this study. It acknowledges my findings, and is currently working on a plan to improve the situation, which should be announced soon.

People who use the word meritocracy as a positive depiction of society are probably unaware of its original satirical and pejorative connotations. It works only if everyone has the opportunity to compete.

If we really believe in meritocracy as the main principle for sorting academics into positions, we must become better at focusing on the subtle and unconscious gender biases enmeshed in our day-to-day work activities. At stake is not just women's participation in science, but also the stature, integrity and legitimacy of a scientific system renowned for its conformity to the meritocratic ideal. Although all researchers in principle should be considered equal, some remain more equal than others. ■

Mathias Wullum Nielsen is a postdoctoral fellow at the Danish Centre for Studies in Research and Research Policy at Aarhus University, Denmark.
e-mail: mwn@ps.au.dk

**DISADVANTAGES
ACCUMULATE
OVER TIME
THROUGH
SELF-
REINFORCING
EFFECTS.**

➔ **NATURE.COM**
Discuss this article
online at:
go.nature.com/mzlydq

RESEARCH HIGHLIGHTS

Selections from the
scientific literature

ZOOLOGY

How the sponge got its skeleton

Sponges build their skeletons using specialized cells that transport and assemble structural beams like construction workers — a novel way of producing a skeleton compared to other animals.

Sponge skeletons are made of rod-like silica structures called spicules that are cemented to rocks and to each other with collagen. To find out how the spicule assembly process works, Noriko Funayama at Kyoto University in Japan and her colleagues studied a freshwater sponge (*Ephydatia fluviatilis*) under a microscope and discovered 'transport cells' that move spicules inside the sponge. The cells then push the spicules through the animals' outer surface to raise them up and attach them together.

This process allows sponges to adopt a huge variety of shapes and sizes, the authors say.

Curr. Biol. <http://doi.org/7sr> (2015)

NEUROSCIENCE

Sound switches on worm cells

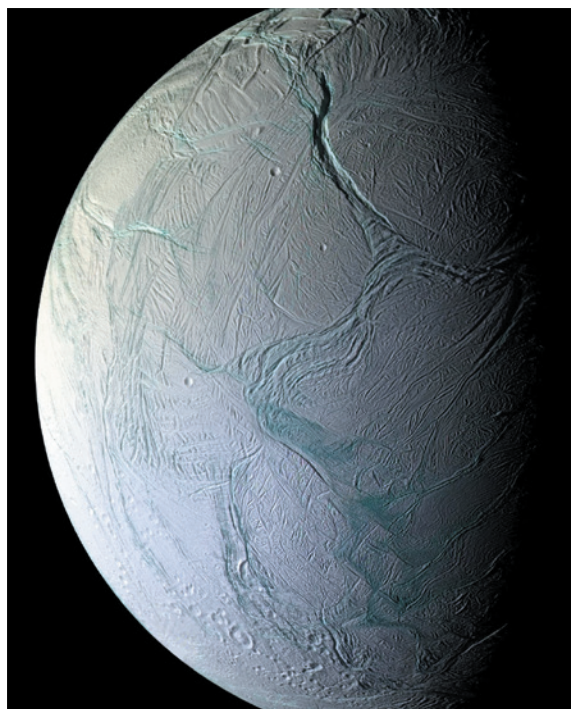
Ultrasound has been used to stimulate individual brain cells in a worm. If the technique works in mice, it could be a less invasive way of studying specific neurons.

Neuroscientists currently implant probes into animal brains to stimulate cells that have been engineered to become sensitive to light. Sreekanth Chalasani at the Salk Institute for Biological Studies in La Jolla, California, and his colleagues instead introduced a pressure-sensitive protein,

TRP-4, into neurons in the nematode *Caenorhabditis elegans*. They then put the worms in a Petri dish that was partially submerged in a water bath and sent a short burst of ultrasound into the dish, delivering mechanical signals to TRP-4 to activate certain neurons.

By adding the TRP-4 protein into neurons with different functions, the researchers were able to make free-crawling worms reverse direction, stop reversing or make more-frequent sharp turns in response to a brief pulse of ultrasound.

Nature Commun. 6, 8264 (2015)



PLANETARY SCIENCE

Global ocean on Enceladus

Beneath an icy crust, Saturn's moon Enceladus (pictured) has an ocean that covers its entire globe.

NASA's Cassini spacecraft measured wobbles in Enceladus's rotation over more than seven years. The data confirm that the crust is moving separately from the rocky core, meaning that there must be a widespread layer of liquid between them, says a team led by Peter Thomas of Cornell University in Ithaca, New York.

Cassini had previously spotted jets of liquid spewing from the moon's surface, and other studies have suggested that Enceladus has an underground sea only near its south pole. This latest finding further highlights how Enceladus could be one of the most likely places for extraterrestrial life.

Icarus <http://doi.org/7rf> (2015)

NEUROSCIENCE

Electric zaps help spinal-cord rehab

Electrically stimulating a damaged spinal cord as part of rehabilitation therapy may enhance improvements in movement.

Steve Perlmuter at the University of Washington in Seattle and his team bruised the spinal cords of rats to partially paralyse the animals' forelimbs. They then used a neural-computer interface connected to the limb muscles and spinal cord to direct an electrical pulse to just below the damaged spinal area whenever the device detected activity in the weakened muscles.

Rats that received pulses for several weeks recovered their ability to reach for and grasp food pellets with their forelimbs to a greater extent than those that did not receive pulses. The stimulated rats maintained their recovery even after the stimulation was stopped, suggesting that it induced lasting changes in the spinal cord. The scientists suggest that the approach might also work in the clinic.

Proc. Natl Acad. Sci. USA <http://doi.org/7q4> (2015)

EVOLUTION

Ancient lung parts found in fish

A fish species found in the Indian Ocean has a vestigial lung, suggesting that its ancestors had working lungs before they shifted to life in deep waters.

The coelacanth fish *Latimeria chalumnae* is descended from ancient coelacanths that lived in shallow waters. Paulo Brito at Rio de Janeiro State University in Brazil and his colleagues studied the fish at different

NASA/JPL/SPACE SCIENCE INST.

stages of development, and found that a lung developed early in the embryo but then slowed its growth as the embryo matured. As the lung shrank in size relative to the growing embryo, a fatty organ that helps fish to control their buoyancy developed.

This is further evidence that ancestral coelacanths could breathe air in shallow waters, and that they lost the use of the lung as it was replaced by the fatty organ — an important adaptation to the deep ocean.

Nature Commun. 6, 8222 (2015)

PLANETARY SCIENCE

A balmy climate on exoplanets?

Certain planets outside our Solar System could have wind patterns that produce habitable climates.

Ludmila Carone at the University of Leuven in Belgium and her team used climate models to investigate atmospheric temperatures and wind patterns on planets with Earth-like atmospheres. The chosen planets closely orbit red-dwarf stars and always face their stars with the same side. The team found 3 possible climates for planets that have orbits of less than 12 days. Two of these climates could potentially host life, because of wind jets that stop the side of the planet exposed to the star getting too hot.

The findings could help to guide the selection of exoplanets for future study, the authors say.

Mon. Not. R. Astron. Soc. 453, 2412–2437 (2015)

PLANT ECOLOGY

Orchid shapes trick male insects

Orchids have adapted the shape of their flowers to attract pollinating wasps.

These flowering plants lure male insect pollinators by producing chemicals that mimic the pheromones of

their female counterparts, but the effect of flower shape on pollinators has been unclear. To look at this, Marinus de Jager and Rod Peakall at the Australian National University in Canberra studied two species of *Chiloglottis* orchids that emit the same pheromone and the two species of *Neozeleboria* wasps that pollinate the flowers. They found that the wasps copulated more frequently and for longer periods of time (pictured) with the orchid that they normally pollinate.

The dimensions and colour of the preferred orchid's callus (the central part of the flower) closely resembled the respective female wasp, and the overall shape of the flower allowed the male wasp to fit better within it.

Funct. Ecol. <http://doi.org/7rd> (2015)

AGRICULTURE

Ecological impact of crops drops

The environmental impact of maize (corn) and cotton crops on US freshwater ecosystems has been decreasing over the past decade, mainly because of the use of genetically modified plants that require less added pesticide.

Sangwon Suh and Yi Yang at the University of California, Santa Barbara, assessed the local environmental impacts of crops, including pollution from direct runoff of fertilizers and pesticides, as well as from processing and transportation. They found that the impact of maize and cotton has decreased by about 50% over the past decade.

However, the impact of soya-bean crops has increased threefold, owing to the

spread of an invasive soya-bean pest and a consequent rise in the use of insecticides. The authors say that further improvements may be more

SOCIAL SELECTION

Popular topics on social media

A call to deal with the data deluge

As the number of biomedical research papers continues its relentless growth, the quality and credibility of science is buckling under the weight of all the data. That is the conclusion of an article in the journal *eLife* that triggered discussion online this week. The piece, which is based on interviews with 20 anonymous US senior scientists, suggests a radical rethinking of the peer-review system to deal with the 'overflow' of data. Erik Müllers, a cell biologist at the Karolinska Institute in Stockholm, summed up the issue on Twitter: "Too many journals, too many researchers, too low quality: Overflow in #science and its implications for trust." But not everyone was

on board. "Is there really an 'overflow' problem in science? I don't think so," tweeted Savraj Grewal, a cell biologist at the University of Calgary in Canada.

eLife 4, e10825 (2015)

➔ **NATURE.COM**
For more on popular papers:
go.nature.com/4seski



difficult, because pests and weeds are beginning to develop resistance to the pesticides produced by the modified crops.

Environ. Res. Lett. 10, 094016 (2015)

GENETICS

How Inuit genomes have adapted

The genomes of indigenous people in Greenland (pictured) show how they have adapted to thousands of years of frigid temperatures and a diet that is rich in fatty seafood.

Rasmus Nielsen at the University of California, Berkeley, and his colleagues analysed the genomes of 191 Inuit people from Greenland and compared

them with genomes from people of European or Han Chinese descent. They found that the Inuit genomes were enriched for genes that convert certain fatty acids in the diet into more biologically active forms, and that counteract the oxidative stress associated with a high-fat diet. The team also discovered a mutation in the Inuit genomes that is linked to the development of brown fat cells, which generate heat.

These mutations seem to date from at least 20,000 years ago, when Inuit ancestors lived around the Bering Strait between Russia and Alaska.

Science 349, 1343–1347 (2015)

➔ **NATURE.COM**
For the latest research published by Nature visit:
www.nature.com/latestresearch



SEVEN DAYS

The news in brief

BUSINESS

Ebola armoury

On 14 September, the US Biomedical Advanced Research and Development Authority (BARDA) awarded pharmaceutical giant Johnson & Johnson US\$28.5 million to manufacture two components of a candidate Ebola vaccine, which is in phase II clinical testing. On 17 September, ZMapp, a cocktail of three monoclonal antibodies that is marketed by LeafBio in San Diego, California, was granted fast-track approval by the US Food and Drug Administration. And on 21 September, BARDA awarded up to \$38 million to Regeneron Pharmaceuticals of Tarrytown, New York, to develop a monoclonal-antibody drug against Ebola.

EVENTS

Refugee scientists

As Germany struggles to cope with an influx of tens of thousands of — mostly Syrian — refugees, German universities and research organizations are wooing talented scientists among them. Migrant scientists and researchers at German institutions can register at www.chance-for-science.de, an online platform created by academics at Leipzig University to connect highly qualified refugees with German scientists. Meanwhile, the Fraunhofer and Max Planck societies have announced a joint initiative to integrate refugee scientists into institutes run by the two organizations.

Malaria milestone

A key target in reducing deaths from malaria has been met, the World Health Organization (WHO) and the UN children's society UNICEF announced on 17 September. Since 2000,

when the UN established malaria reduction as one of its Millennium Development Goals, malaria incidence has dropped by 37% worldwide. Malaria deaths have fallen by 60%, saving 6.2 million lives. But the agencies warn that the battle is not over: more than 200 million new cases have occurred in 2015 alone. The WHO's latest roadmap calls for a further 90% reduction in malaria cases and deaths by 2030.

The end of STAP

Controversial claims that embryonic-like cells could be produced by exposing adult cells to acid have been laid to rest in two papers this week (A. De Los Angeles *et al.* *Nature* <http://dx.doi.org/10.1038/nature15513> (2015) and D. Konno *et*

al. *Nature* <http://dx.doi.org/10.1038/nature15366>; 2015). Two 2014 *Nature* papers that made the claims about 'stimulus-triggered acquisition of pluripotency' (STAP) cells were retracted. The follow-up papers report that 7 labs failed in 133 attempts to produce STAP cells. Genetic tests found several cases in which purported STAP cells created at the RIKEN Center for Developmental Biology in Kobe, Japan, were matches for pre-existing embryonic stem cells.

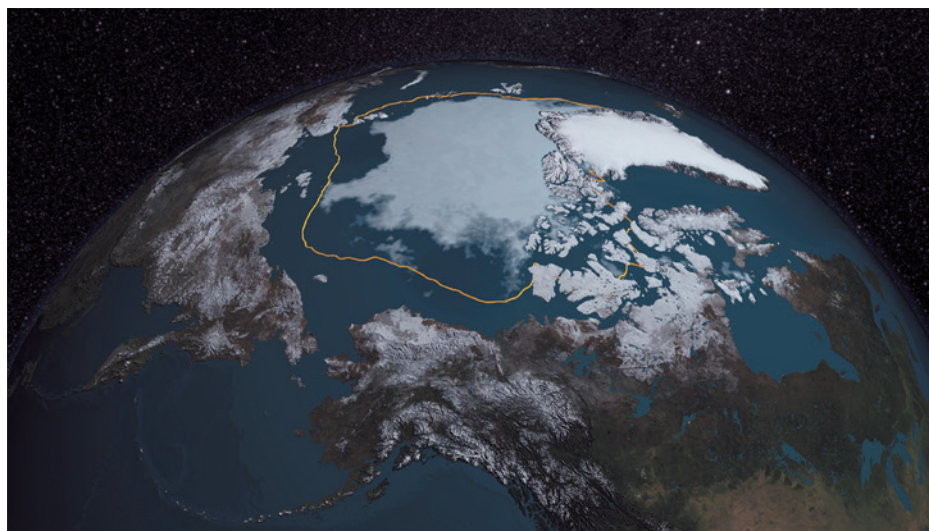
Orion approval

NASA has approved the next stage of a project to build a spacecraft that could send astronauts to Mars and beyond. The 16 September approval came after a review of the project, dubbed Orion. NASA

has committed US\$6.77 billion to Orion. The decision follows a successful uncrewed flight of the spacecraft in December 2014. NASA also committed to a test flight with astronauts no later than April 2023.

Car emissions bust

The US Environmental Protection Agency (EPA) began investigating German car maker Volkswagen on 18 September for designing vehicles that can bypass US emissions standards. Volkswagen officials admitted installing software in some models that switches on all the emissions controls only when it detects that a car is undergoing an emissions test. During normal driving conditions, the vehicles emit up to 40 times more nitrogen oxides than allowed by law, according to



NASA/GODDARD SCIENTIFIC VISUALIZATION STUDIO

Another low for Arctic sea ice

This summer, Arctic sea ice reached its fourth-lowest coverage since satellite records began in 1979, the US National Snow and Ice Data Center (NSIDC) in Boulder, Colorado, announced on 15 September. The minimum surface area of Arctic ice for 2015 is 34% less than the average

the EPA. The agency says that the software was installed on at least 482,000 diesel vehicles sold in the United States.

FACILITIES

LIGO is 'go'

Advanced LIGO (Laser Interferometer Gravitational-Wave Observatory) officially began gathering data on 18 September. LIGO's twin instruments, in Louisiana and Washington state, each have two 4-kilometre arms, and represent a US\$200-million overhaul of the initial LIGO, which attempted to detect gravitational waves in the 2000s. These space-time ripples are one of the major predictions of Einstein's general theory of relativity that have still to be observed directly. On the same day, Italy and France agreed to extend their collaboration on Virgo, LIGO's European counterpart, for a further five years beyond 2020.

FUNDING

Beaver genome bid

Genomics researchers are hoping that sports fans at Oregon State University in Corvallis will fund an effort to sequence the genome of the college mascot: the North American beaver (*Castor canadensis*; pictured). The Beaver Genome Project launched on 16 September



and hopes to raise US\$30,000 by 30 October (go.nature.com/cjaizw). The funds will pay for genome sequencing of Filbert, a four-year-old beaver born and raised at Oregon Zoo in Portland. Researchers hope to gain insight into the rodent's ability to digest wood and its complex dam-building behaviour.

RESEARCH

CRISPR request

The UK Human Fertilisation and Embryology Authority (HFEA), which regulates embryo research, is considering the country's first application for work that would involve editing the genomes of human embryos. Genome editing is illegal for treatment in the United Kingdom, but is possible for research under licence from the HFEA. The application to use the CRISPR/Cas9 technique comes from Kathy Niakan, a researcher at the Francis Crick Institute in London. Earlier this year, a

Chinese team used the same method to edit the genomes of human embryos.

PEOPLE

Food and drug chief

US President Barack Obama has nominated cardiologist Robert Califf to become the next head of the US Food and Drug Administration, the White House announced on 15 September. Califf, a clinical-trial specialist, is currently taking a leave of absence from Duke University in Durham, North Carolina, while serving as the deputy commissioner of the agency. The US Senate must confirm his nomination before Califf can take the position.

Vatican stargazer

Pope Francis announced on 18 September that US planetary scientist Guy Consolmagno has been appointed director of the Vatican Observatory. A Jesuit brother, science-fiction fan and co-author of the book *Would You Baptize an Extraterrestrial?*, Consolmagno studies

COMING UP

27 SEPTEMBER –
2 OCTOBER

Europlanet, the European Planetary Science Congress, is held in Nantes, France.
www.epsc2015.eu

28 SEPTEMBER

The Indian Space Research Organisation's astronomy satellite, ASTROSAT, is set to launch from the Satish Dhawan Space Centre at Sriharikota.
<http://astrosat.iucaa.in>

28 SEPTEMBER –
2 OCTOBER

The future quantum world is discussed at the 5th international quantum cryptography conference in Tokyo.
<http://2015.qcrypt.net>

meteorites and asteroids and curates the Vatican's meteorite collection. He replaces José Funes, who had been director since 2006. The observatory is based at Castel Gandolfo, outside Rome, and runs a telescope in Tucson, Arizona.

Australian science

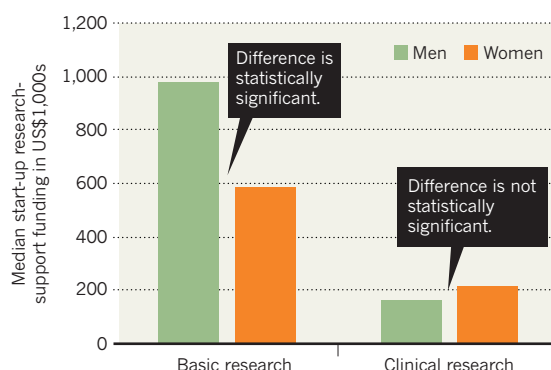
Christopher Pyne was appointed Australia's Minister for Industry, Innovation and Science on 20 September. Prime Minister Malcolm Turnbull shuffled the cabinet within days of his 14 September coup over former prime minister Tony Abbott, whose tenure was marked by drastic cuts to research funding. Earlier this year, as education and training minister, Pyne was involved in a controversial decision to award funding to the University of Western Australia in Perth for a research centre linked to climate-policy sceptic Bjorn Lomborg. The university ultimately rejected the funding.

TREND WATCH

In New England, male researchers in basic biomedical science receive 68% more lab start-up funding than do female applicants, shows a 15 September study (R. Sege *et al.* *J. Am. Med. Assoc.* **314**, 1175–1177; 2015). A team analysed grant applications in 2012–14 to the Medical Foundation Division of Health Resources in Action (HRA) in Boston, Massachusetts, to calculate institutional support given to 219 applicants of 2 HRA programmes. The disparity was not explained by experience, degree or host-institution wealth.

GENDER GAP IN START-UP SUPPORT

Male and female junior researchers in the New England region of the United States fare very differently in the start-up funding they get — which includes salary, research technicians, equipment and supplies.



NEWS IN FOCUS

SUSTAINABILITY Scientists call for sharper focus in UN anti-poverty push **p.434**

BRAIN STIMULATION Studies in children raise hopes and fears for the technology **p.436**

ASTROSAT Indian space observatory excites astronomers globally **p.438**



PRIVACY Researchers seek safe ways to mine confidential data **p.440**

EPH/SEDAT SUNA/CORBIS



Syrian refugees wait for a bus in Istanbul, Turkey: many young people are missing out on higher education as a result of conflict in the Middle East.

MIDDLE EAST

Lost generation looms as refugees miss university

Educational void risks hampering reconstruction in Middle East.

BY DECLAN BUTLER

Human-rights organizations are calling on universities and governments worldwide to invest more in the education of the hundreds of thousands of student refugees who are fleeing war-torn regions of the Middle East.

They warn that the countries in conflict risk losing a future generation of scientists, engineers, physicians, teachers and leaders — and that university-aged refugees who have found shelter elsewhere represent a crucial

opportunity to reverse some of the lost intellectual capital. “Each scholar and student that we lose now deepens the challenge of restoring the region when the violence eventually subsides,” says Robert Quinn, executive director of the Scholars at Risk Network, a human-rights group headquartered in New York City.

Quinn also cautions that allowing an educational void to develop in the Middle East could create a fertile recruiting environment for radical militias and terrorists. “It is deeply in the interest of Europe and the West to protect and invest in the intellectual capital of the region,” he

says. “The failure to invest massively is foolishly shortsighted.”

Conflicts in Syria, Iraq and Yemen, as well as in Libya and other North African countries, have led to a record number of refugees. By the end of 2014, 60 million people worldwide were seeking refuge either in safer parts of their countries or abroad, according to the Office of the United Nations High Commissioner for Refugees. That is the highest number ever recorded, and almost double the 37.5 million displaced individuals a decade earlier.

Syria, which had a population of nearly ▶

► 21 million before the ongoing conflict there began four years ago, has produced the most refugees, with 7.6 million people displaced internally and a further 4 million forced to flee the country. Around 10% of those people are of university age, estimates James King, who is a senior researcher at the Scholar Rescue Fund, part of the Institute of International Education (IIE), a non-profit, educational-exchange organization in New York City.

Yet the university system in Syria has all but collapsed, and few of the young people who have left the country are receiving higher education. Of those refugees who fled abroad, most have found temporary shelter in neighbouring countries — Turkey is hosting some 1.8 million, Lebanon 1.2 million and Jordan 630,000 — but only around 5% of the university-aged refugees in these countries are enrolled at local institutions, according to a March report funded by the European Commission (see go.nature.com/9ljpb1).

Before the conflict began, 26% of young adults in Syria were receiving tertiary education. That leaves hundreds of thousands of people who would normally be attending university going without.

Even when universities in the refugees' host countries have capacity — and this in itself is an issue; Turkey, for example, is struggling to accommodate all of its own eligible and interested students — there are a string of further impediments to enrolment. Many students have fled without documents, says King, including records of past academic credentials. Other issues are financial and material hardships, which can force young adults to work, leaving them no time for education. In Turkey, where just 1% of Syrian refugees aged 18–24 have found university places, language

difficulties are a big problem.

Scholarships are available. The IIE-led Syria Consortium for Higher Education in Crisis, a network of higher-education institutions worldwide that was created in 2012, has provided US\$4.5 million to support 333 Syrian students, including 158 scholarships to attend universities in Western countries. At least 20 similar initiatives also offer scholarships to institutions across the globe. However, demand far outstrips supply: these combined efforts have been able to provide only around 7,000 students with

“Education is the orphan of all these crises. People are so concerned about food, water, shelter and other basics.”

some form of tertiary education.

Allan Goodman, president and chief executive of the IIE, notes the sheer scale of the crisis. “No organization or country is set up to deal with it,” he says, “The only thing we can do is — one by one, family by family, scholar by scholar, student by student — try to help individuals.”

He also says that humanitarian efforts have tended to focus on saving lives and relieving misery among those fleeing conflict. “Education is the orphan of all these crises,” he says. “People are so concerned about food, water, shelter and other basics, and we haven’t thought enough about education.” The 1.5% of global humanitarian aid that goes to education, meanwhile, is spent largely on primary and secondary schooling, not higher education, which traditionally has been seen as a luxury.

There are signs that attitudes are changing. In May, the European Union’s trust fund for the Syrian crisis committed €12 million (US\$14.5 million) to assist 20,000 Syrian

refugees in obtaining higher education through scholarships and other means. As the European Commission report notes, however, scholarships cannot meet the enormous need, which would amount to billions, not millions, of euros.

It would be more cost effective to provide direct financial aid to universities in the countries with the most Syrian refugees, the report states. And various organizations, including the UN children’s charity, UNICEF, are exploring whether the massive open online courses (MOOCs) now offered by some top universities could also help. By using recorded lectures and social-networking-style communication, for example, MOOCs are intended to democratize access to a world-class education. But they are largely untested in a refugee situation, Goodman says, and most students still want a diploma accredited by a ministry of education. “A MOOC from Stanford or MIT isn’t the same,” he says. “The most durable situations are those that seek to integrate students into national university systems.”

Clearly, a long-term solution will require enormous investment and much greater involvement by higher-education institutions worldwide, Quinn says. Next month, the IIE and other organizations will hold a two-day workshop in Istanbul, Turkey, aimed at better coordinating efforts and exploring fresh approaches to scaling up access.

The challenge is great — not least because the conflicts seem set to get worse before they get better. “But this must be measured against the costs of not doing it,” says Quinn. “If we invest over the next five or ten years in educating and strengthening as many Middle East citizens and children as possible, we will have planted the seeds of a transformed region and much brighter future for the world.” ■

SUSTAINABLE AID

UN sets out next development goals

Scientists call for sharper focus in anti-poverty push.

BY JEFF TOLLEFSON

On 25 September, Pope Francis will address the United Nations just before a three-day meeting that will set the agenda for international development efforts over the next 15 years. At the Sustainable Development Summit in New York, global leaders will adopt 17 goals that are meant to improve the lives of the world’s

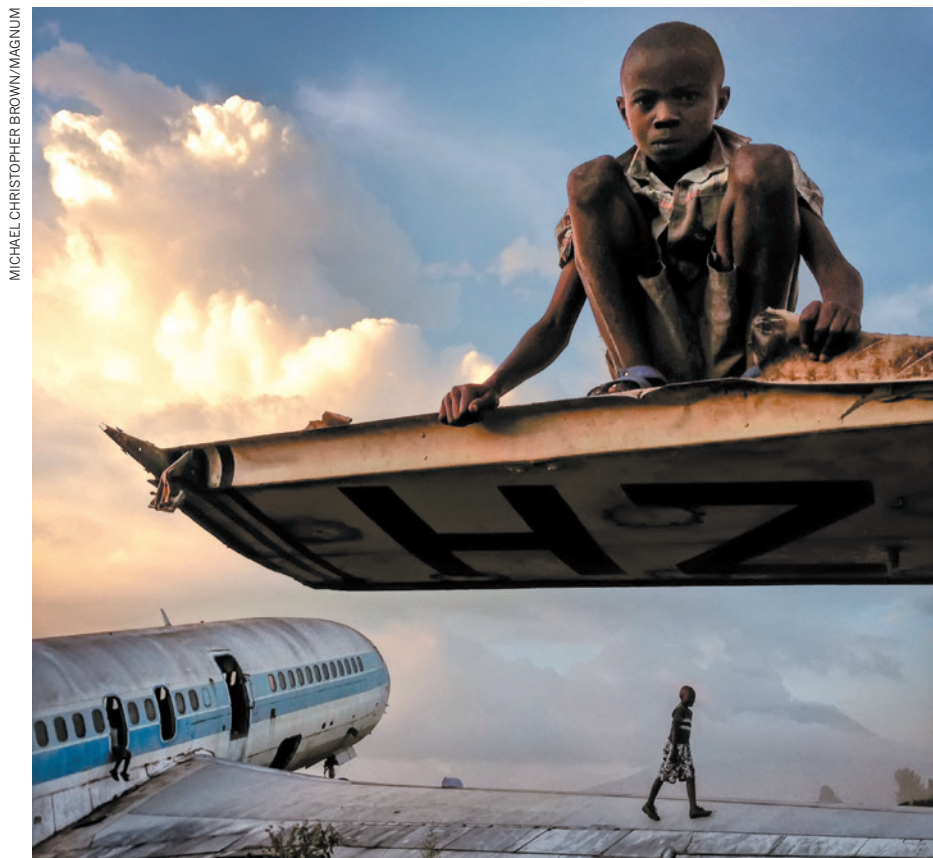
poorest people by 2030, without jeopardizing the health of the planet.

Ambitious and broad, these Sustainable Development Goals (SDGs) would, if met, greatly improve human welfare. But some experts fear that the goals are too numerous and vague to have practical value. “I’m a little worried that there are too many of them,” says Steven Radelet, director of the Global Human Development Program at Georgetown

University in Washington DC. “They may fall prey to the old adage that if everything is a priority, then nothing is a priority.”

First on the list: “End poverty in all its forms everywhere”. Second is to end hunger and achieve food security while improving nutrition and promoting sustainable agriculture. The list goes on to address fundamental issues such as education, gender equality and access to water and basic sanitation services. It also calls for economic growth, environmental conservation and clean energy for all people, while urging action to combat climate change. The goals are supplemented by 169 specific targets that are meant to clarify the work that needs to be done.

Under discussion since 2012, the SDGs replace the expiring Millennium Development Goals, which the UN adopted in 2000. Those eight objectives called, among other things, for halving extreme poverty, reducing mortality among children under five by two-thirds and instituting universal primary



The UN goal of ending poverty would help these street children in the Democratic Republic of the Congo.

education, all by 2015. Although the world has made considerable advances in many of these areas, it has been debated how much impact the goals themselves have had. Much of the progress in alleviating poverty over the past two decades, for example, has come from rapid economic development in southeast Asia and China.

Moving forward, the challenge for governments will be to invest limited resources effectively and track progress towards the goals. The UN is discussing how to structure a progress-assessment system based on a list of measurable indicators that it is developing for the targets. As it stands, the list is too long for governments seeking to measure their progress, says Mark Stafford Smith, a researcher at Australia's Commonwealth Scientific and Industrial Research Organisation in Canberra.

It will be up to the scientific community to identify simpler indicators and policies that will promote progress, says Stafford Smith, who chairs the scientific committee of Future Earth, an international clearing house for sustainability research. Replacing firewood with a more sustainable fuel source, for instance, would boost air quality and therefore improve human health (goal 3), while reducing the impact on local ecosystems (goal 15). And by reducing time spent foraging for fuel, it would free children to go to school (goal 4) and empower women to

contribute to economic growth by earning money (goals 5 and 8).

Researchers also want to find ways to prevent conflicts between goals. For example, without advances in efficiency and a shift towards renewable energy, the expansion of access to modern energy sources (goal 7) would interfere with the goal of keeping global warming in check (goal 13). “People have piled everything in there, but the research community can focus on a much smaller set of integrated goals.”

Stafford Smith. “If we don’t do that, then we will find that these potential conflicts become real ones.”

Through a project called The World in 2050, researchers who have used computer models to explore the socio-economic implications of climate change are leading an analysis to identify policy scenarios that can assure that the goals are met over the next few decades.

One of the institutions leading the effort is the International Institute for Applied Systems Analysis in Laxenburg, Austria. Its deputy director-general, Nebojsa Nakicenovic, says: “The idea is to understand how one can achieve all of those goals together.” ■

Brain stimulation in children spurs hope — and concern

Treatment of developing brains offers greater scope for improvement but also intensifies risks.



Studies aimed at enhancing learning in children are generating controversy.

BY LINDA GEDDES

Jack struggled in regular school. Diagnosed with dyslexia and the mathematical equivalent, dyscalculia, as well as the movement disorder dyspraxia, Jack (not his real name) often misbehaved and played the class clown. So the boy's parents were relieved when he was offered a place at Fairley House in London, which specializes in helping children with learning difficulties. Fairley is also possibly the first school in the world to have offered pupils the chance to undergo electrical brain stimulation.

The stimulation was done as part of an experiment in which twelve eight- to ten-year-olds,

including Jack, wore an electrode-equipped cap while they played a video game. Neuroscientist Roi Cohen Kadosh of the University of Oxford, UK, who led the pilot study in 2013, is one of a handful of researchers across the world who are investigating whether small, specific areas of a child's brain can be safely stimulated to overcome learning difficulties. "It would be great to be able to understand how to deliver effective doses of brain stimulation to kids' brains, so that we can get ahead of developmental conditions before they really start to hold children back in their learning," says psychologist Nick Davis of Swansea University, UK.

The idea of using magnets or electric currents

to treat psychiatric or learning disorders — or just to enhance cognition — has generated a flurry of excitement over the past ten years. The technique is thought to work by activating neural circuits or by making it easier for neurons to fire. The research is still in its infancy, but at least 10,000 adults have undergone such stimulation, and it seems to be safe — at least in the short term. One version of the technology, called transcranial magnetic stimulation (TMS), has been approved by the US Food and Drug Administration to treat migraine and depression in adults.

Interest is growing, however, in whether such technologies might have even greater benefits in children. Particularly promising is TMS's cheaper and more-portable cousin, transcranial direct-current stimulation (TDCS).

Researchers say that the stimulation effects are likely to penetrate deeper in children because their skulls are thinner than adults', and might have more of an impact in brains that are still growing. However, the same factors that intensify the potential benefits are also cause for concern. "It's like when you build a house: if you think things are going wrong, it's much easier to fix things at the beginning rather than later on, but it's also much easier to ruin them," says Cohen Kadosh. "We don't know how electrical stimulation interacts with the developing brain."

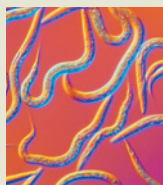
Cohen Kadosh also worries about abuse of the technology. Although devices prescribed for medical treatments must meet certain safety standards, there are currently no laws in either Europe or the United States to regulate the use of TDCS in people merely hoping to enhance their cognition, and companies now sell the TDCS headsets online. So parents, say, might feel tempted to try to boost the cognitive abilities of their children outside the controlled conditions of a lab. After weighing up the pros and

KLAUS VEDFELT/GETTY



**MORE
ONLINE**

TOP STORY



Chinese scientists row over long-sought protein that senses magnetism
go.nature.com/eln84c

MORE NEWS

- Camouflaged nanoparticles evade immune system go.nature.com/buzdmv
- Dramatic rise seen in antibiotic use go.nature.com/kiove2
- Use of personalized cancer drugs runs ahead of the science go.nature.com/xujlgw

NATURE PODCAST



Malaria interventions, using private data in research and twisted neutrons
nature.com/nature/podcast

cons, however, Cohen Kadosh decided to approach Fairley House about doing a trial. He also had to seek ethics approval, which he received. "We were very worried about putting brain stimulation into place, because as a school we knew nothing about it, but we were reassured about the ethics and safety," says Jenny Lim, an occupational therapist who works with children at the school.

LEARNING ENHANCER

The study follows on from one in which Cohen Kadosh showed that a variant of TDCS called transcranial random-noise stimulation (TRNS) could boost mathematical ability in adults (A. Snowball *et al.* *Curr. Biol.* **23**, 987–992; 2013).

In the Fairley House study, his team gave the 12 children with mathematical learning difficulties nine 20-minute training sessions. Half of the volunteers received TRNS, targeted at the brain area responsible for processes such as planning and abstract reasoning; the other half wore a TRNS cap but did not receive any stimulation. TRNS is thought to work by modulating brain signals during learning: in this case, the children moved their bodies from side to side to guide a ball on a screen to land at a certain point on a number line, with the difficulty increasing as they progressed.

The children who received stimulation showed greater progress in performance than did the controls — reaching level 17 on average, compared with level 14 — as well as significant improvements in general mathematics test scores. Cohen Kadosh presented the analysis at the British Association for Psychopharmacology meeting in Bristol in late July and has submitted the results for publication. He now plans to further this line of research.

But neuroscientist Vincent Walsh at University College London's Institute of Cognitive Neuroscience thinks that studies of brain stimulation in children are premature. The benefits observed in young adults are not always seen in older people, he says, and many electrical-stimulation results have yet to be replicated. "There is simply no sound scientific basis for extending such poor work to children," he says.

Davis, by contrast, thinks that such experiments are justified, but is concerned about the trend to use the techniques outside formal studies. He estimates that at least 1,000 children around the world have received some kind of brain stimulation as part of clinical studies, and expects more in future. He stresses the importance of publishing the results of any work done in children. "I would urge all scientists to share their results when children and young people are given brain stimulation, to allow other scientists to learn from 'failed' trials and to adapt the protocols if needed." ■

POLITICS

Science vies for notice in Canadian election

Current government has cut funding and limited researchers' influence over policy, critics say.

BY NICOLA JONES

Canadians will head to the polls on 19 October in a federal election that many scientists hope will mark a turning point after years of declining research budgets and allegations of government censorship.

Prime Minister Stephen Harper, who has been in office since 2006, now finds his right-leaning Conservative party in a tight three-way race with the left-leaning New Democratic Party (NDP) and the middle-left Liberals. Although science has not emerged as a top issue during the campaign, researchers are fighting to make their concerns heard.

In an unprecedented move, the Professional Institute of the Public Service of Canada — a union in Ottawa that represents more than 57,000 government scientists and other professionals — is campaigning in the federal race. "Here's how we do things in the Harper government," declares one of the union's radio advertisements. "We muzzle scientists, we cut research and we ignore anyone who doesn't tell us what we want to hear."

The group estimates that the Harper administration has eliminated jobs for some 2,500 scientists. And the government's own data show that Canada's ranking for research and development spending dropped from 16th among 41 comparable nations in 2006 to 23rd in 2011 (the most recent year for which government figures are available). Harper has also been accused of limiting government scientists' ability to communicate with the press and public; Canada's information commissioner promised to investigate this in 2013, but has not yet released any findings.

"The Harper government has had complete disdain for federal government science," says Peter Wells, a marine biologist at Dalhousie University in Halifax, Nova Scotia.

Kai Chan, an ecologist at the University of British Columbia in Vancouver, is similarly glum. "I have been continually surprised by how bad it has gotten in Canada," says Chan, who co-founded a group called 'scienceinpolicy' in the United States during George Bush's presidency. "It's worse than I could have imagined, having closely scrutinized what I thought was the worst in North America."

Although the Conservative party has done little to address such criticisms, members of the two opposition parties have called for a stronger role for science in government. Dozens of NDP and Liberal candidates for parliament have declared their support for evidence-based decision-making by signing a 'science pledge' developed by Evidence for Democracy, a non-profit science-advocacy group in Ottawa.

These issues were not discussed at a debate in Calgary on 17 September that pitted Harper against NDP leader Thomas Mulcair and Liberal leader Justin Trudeau, and focused on the state of Canada's economy. But concerns about the condition of Canadian science

have nevertheless influenced party platforms. The Liberal Party has made scientific integrity part of its election campaign, proposing the creation of a central public portal for disseminating government-funded research. The party is seeking to appoint a chief science officer to ensure the free flow of information.

By contrast, Harper's government phased out the position of national science adviser in 2007–08, replacing it with the Science, Technology and Innovation Council (STIC), a body that has drawn criticism from scientists for operating behind closed doors. STIC reviews issues at the request of the federal government, and many of its reports are confidential. "It doesn't even have many scientists on it," says Graham Bell, a biologist at McGill University in Montreal and president of the Royal Society of Canada, who would like to see the science adviser's post re-established.

Similarly, the NDP has called for a parliamentary science officer, a position that would be independent of the majority party or a coalition leading the government. Katie Gibbs, executive director of Evidence for Democracy, says that Canada could benefit from either a science adviser or a parliamentary science officer. "They're different visions," she says. "You could easily have both." ■

public portal for disseminating government-funded research. The party is seeking to appoint a chief science officer to ensure the free flow of information.

By contrast, Harper's government phased out the position of national science adviser in 2007–08, replacing it with the Science, Technology and Innovation Council (STIC), a body that has drawn criticism from scientists for operating behind closed doors. STIC reviews issues at the request of the federal government, and many of its reports are confidential. "It doesn't even have many scientists on it," says Graham Bell, a biologist at McGill University in Montreal and president of the Royal Society of Canada, who would like to see the science adviser's post re-established.

Similarly, the NDP has called for a parliamentary science officer, a position that would be independent of the majority party or a coalition leading the government.

Katie Gibbs, executive director of Evidence for Democracy, says that Canada could benefit from either a science adviser or a parliamentary science officer. "They're different visions," she says. "You could easily have both." ■

BIOMEDICINE

Scripps fills top posts

Duo to focus on finances after a failed merger.

BY ERIKA CHECK HAYDEN

A geneticist and a chemist will co-lead the Scripps Research Institute in La Jolla, California, the biomedical research organization announced on 18 September. The institute said that Scripps chemist Peter Schultz will take over as chief executive and vice-chair, while molecular biologist Steve Kay, currently at the University of Southern California (USC) in Los Angeles, will assume the institute's presidency.

The appointments are the latest in a series of leadership changes at the research institute. Over the past decade, public funding from the US National Institutes of Health has flattened, competition for philanthropic donations has intensified and pharmaceutical companies have shifted away from providing unrestricted funds for basic research. This has meant that Scripps and other independent research organizations have struggled to stay afloat.

"The broader significance is this worldwide need to change the model for drug discovery, and to show that an effective bench-to-bedside model can be created within the not-for-profit sector," says Kay, who is dean of the Dornsife College of Letters, Arts and Sciences at USC.

Kay and Schultz's appointments come just over a year after the failure of a bid by USC to merge with Scripps. In July 2014, Scripps president and chief executive Michael Marletta departed the institute after a faculty revolt against the merger deal. The following month, James Paulson, head of the institute's cell- and molecular-biology department, was named acting president and chief executive. When the merger fell through, Scripps was said to be running an operating deficit of US\$21 million.

"Looking forward, I think many scientists realize that NIH funding is a good thing if you have it, but it's not sustainable," says organic chemist Phil Baran, who was on the search committee that selected Kay and Schultz. "What is stable are endowments, which you build by having products that give you proceeds, and by philanthropy. You get philanthropy by doing the best science, so that's why there is such frenzied competition for the brightest minds."

Kay, who was previously at Scripps from 1996 to 2007, hopes that his experience with the institute will ease any tensions about his arrival from USC. ■



Multiple devices will allow ASTROSAT to access more wavelengths than most other telescopes.

ASTRONOMY

Indian telescope set for global stardom

ASTROSAT will extend the capabilities of existing US and European observatories, and boost Indian research.

BY T. V. PADMA

A satellite is about to bring India international acclaim, at least in astronomy circles. On 28 September, ASTROSAT, the country's first space observatory dedicated to science, will take to the skies.

As well as boosting the activities of Indian astronomers — who are abuzz with excitement — the satellite is expected to benefit researchers all over the world. Designed to orbit Earth for five years, it has capabilities not offered by existing space telescopes.

"It is a notable and fantastic step forward for Indian astronomy, and has broad implications for astronomers everywhere," says Henry Yang, a mechanical engineer at the University of California, Santa Barbara, who is chair of the board for the Thirty Meter Telescope (TMT) project, an observatory planned for Mauna Kea, Hawaii.

India has had ground-based telescopes for decades, including the Giant Metrewave Radio Telescope near Pune and the Indian Astronomical Observatory in the Himalayan cold desert of Ladakh. But although these can detect radio waves and infrared radiation,

which easily penetrate Earth's atmosphere, they cannot monitor higher frequencies that the atmosphere tends to block — most ultraviolet light, for example, and all X-rays and γ-rays. Without a space telescope of their own, Indian scientists have had to rely on ones operated by NASA and the European Space Agency (ESA) to study such radiation bands, which carry information about exotic neutron stars, newly born or exploding stars and the spiralling hot gases around black holes.

"Often, as we do not know the exact specifics of the telescope design, we are not able to tune our research proposals accordingly," says Varun Bhalerao, an astrophysicist at India's Inter-University Centre for Astronomy and Astrophysics (IUCAA) in Pune.

Indian astronomers have long been at a disadvantage for X-ray and ultraviolet studies, says Somak Raychaudhury, who is the director of the IUCAA and has been involved with ASTROSAT since its inception. Orbiting 650 kilometres above Earth, ASTROSAT will collect data on this portion of the light spectrum, giving Indian scientists faster — and guaranteed — access to the information. They will also have privileged access. "Everybody,

senior or junior scientists, is talking about studies they can now propose,” adds Bhalerao, who is excited about studying neutron stars from India without having to wait for international support. Bhalerao has been studying these stellar objects using high-energy X-ray wavelengths with NASA’s Nuclear Spectroscopic Array (NuSTAR) at the California Institute of Technology in Pasadena, and is looking forward to extending that study to the lower-energy X-ray and ultraviolet bands that will be available through ASTROSAT.

With five instruments, or ‘payloads’, tuned to detect different types of light, ASTROSAT will observe a wider variety of wavelengths than most other satellites, from visible light to the ultraviolet and X-ray bands. Mylswamy Annadurai, director of the Indian Space Research Organisation’s Satellite Centre in Bangalore, calls this “the strength and uniqueness of ASTROSAT”. Black holes, galaxy clusters and other celestial objects can blaze with different wavelengths as different events occur. “When all payloads are combined, ASTROSAT gives a coverage which no other observatory has achieved till now,” he says.

For some researchers, the satellite’s X-ray detection capability will fill the gap left when NASA’s Rossi X-ray Timing Explorer satellite died in 2012, after 16 years of operations. Like Rossi, ASTROSAT will look regularly at large

areas of the sky, enabling it to track simultaneously a large number of X-ray sources that change with time, says Randall Smith, an astronomer at the Harvard–Smithsonian Center for Astrophysics in Cambridge, Massachusetts. By contrast, the X-ray telescopes currently in space generally focus on studying individual objects in great detail.

ASTROSAT’s X-ray detectors can also cope with very bright objects that would saturate those on other satellites such as NASA’s

Chandra X-ray Observatory or ESA’s X-ray Multi-Mirror (XMM-Newton) mission. According to Andrew Fabian at the University of Cambridge’s Institute of Astronomy in the United Kingdom, this capability will make ASTROSAT “invaluable” for alerting the international community to short-lived bursts of X-rays — a key indicator that something new is happening in space. ■

Additional reporting by Alexandra Witze.

CORRECTIONS

The Editorial ‘Too close for comfort?’ (*Nature* **525**, 289; 2015) incorrectly stated: “In his defence, Folta argued that the money supported only travel and outreach, not research, and he was therefore under no obligation to disclose it”. Folta did not say this. He said that he had complied with his university’s disclosure rules. The News Feature ‘Why interdisciplinary research matters’ (*Nature* **525**, 305; 2015) incorrectly affiliated Rebekah Brown with Monash University’s Water for Liveability centre — she is director of the Monash Sustainability Institute. The News story ‘Africa braced for snakebite crisis’ (*Nature* **525**, 299; 2015) wrongly described snakes

as ‘poisonous’ instead of ‘venomous’. And the News Feature ‘Team science’ (*Nature* **525**, 308–311; 2015) gave the wrong authors for the report *Evaluating Interdisciplinary Research*. It was written by Veronica Strang and Tom McLeish.

CLARIFICATION

The Editorial ‘Protection priority’ (*Nature* **525**, 290; 2015) made reference to the fact that the mice in the experiments showed no visible sign of distress. That statement referred only to the animals for which the data were not withdrawn. The committee did not comment on whether or not the animals in the withdrawn experiments showed distress.



THE BIG PEEK

The data contained in tax returns, health and welfare records could be a gold mine for scientists — but only if they can protect people's privacy.

BY ERIKA CHECK HAYDEN

In 2011, six US economists tackled a question at the heart of education policy: how much does great teaching help children in the long run? They started with the records of more than 11,500 Tennessee schoolchildren who, as part of an experiment in the 1980s, had been randomly assigned to high- and average-quality teachers between the ages of five and eight. Then they gauged the children's earnings as adults from federal tax returns filed in the

BARTHOLOMEW COOKE/TRUNK ARCHIVE

2000s. The analysis¹ showed that the benefits of a good early education last for decades: each year of better teaching in childhood boosted an individual's annual earnings by some 3.5% on average. Other data showed the same individuals besting their peers on measures such as university attendance, retirement savings, marriage rates and home ownership.

The economists' work was widely hailed in education-policy circles, and US President Barack Obama cited it in his 2012 State of the Union address when he called for more investment in teacher training.

But for many social scientists, the most impressive thing was that the authors had been able to examine US federal tax returns: a closely guarded data set that was then available to researchers only with tight restrictions. This has made the study an emblem for both the challenges and the enormous potential power of 'administrative data' — information collected during routine provision of services, including tax returns, records of welfare benefits, data on visits to doctors and hospitals, and criminal records. Unlike Internet searches, social-media posts and the rest of the digital trails that people establish in their daily lives, administrative data cover entire populations with minimal self-selection effects: in the US census, for example, everyone sampled is required by law to respond and tell the truth.

This puts administrative data sets at the frontier of social science, says John Friedman, an economist at Brown University in Providence, Rhode Island, and one of the lead authors of the education study¹. "They allow researchers to not just get at old questions in a new way," he says, "but to come at problems that were completely impossible before."

PROBING THE POPULATION

In the past few years, administrative data have been used to investigate issues ranging from the side effects of vaccines² to the lasting impact of a child's neighbourhood on his or her ability to earn and prosper as an adult³. Proponents say that these rich information sources could greatly improve how governments measure the effectiveness of social programmes such as providing stipends to help families move to more resource-rich neighbourhoods.

But there is also concern that the rush to use these data could pose new threats to citizens' privacy. "The types of protections that we're used to thinking about have been based on the twin pillars of anonymity and informed consent, and neither of those hold in this new world," says Julia Lane, an economist at New York University. In 2013, for instance, researchers showed that they could uncover the identities of supposedly anonymous participants in a genetic study simply

by cross-referencing their data with publicly available genealogical information (see *Nature* **497**, 172–174; 2013).

Many people are looking for ways to address these concerns without inhibiting research. Suggested solutions include policy measures, such as an international code of conduct for data privacy, and technical methods that allow the use of the data while protecting privacy. Crucially, notes Lane, although preserving

foundations and universities created the Institute for Research on Innovation and Science at the University of Michigan in Ann Arbor to combine university and government data and measure the impact of research spending on economic outcomes. And in July, the US House of Representatives passed a bipartisan bill to study whether the federal government should provide a central clearing house of statistical administrative data.

Yet vast swathes of administrative data are still inaccessible, says George Alter, director of the Inter-university Consortium for Political and Social Research based at the University of Michigan, which serves as a data repository for

approximately 760 institutions. "Health systems, social-welfare systems, financial transactions, business records — those things are just not available in most cases because of privacy concerns," says Alter. "This is a big drag on research."

UNSOUGHT INTIMACY

Feeding those concerns is the rising public unease about online privacy in general. Private companies known as data brokers operate on a vast scale, collecting and selling information about Internet searches, online purchases and other data streams that can be combined to draw surprisingly intimate conclusions. In one famous example, the US retailer Target inferred that a teenage girl was pregnant based on her purchases there, and it began sending her coupons for baby products; her father was alerted to his impending grandchild only when the coupons arrived at the family's home. In a 2014 study⁴ of data brokers, the US Federal Trade Commission pointed out the many ways in which this kind of information could harm consumers. People who buy products such as blood-sugar monitors, for instance, might be placed into a 'diabetes risk' marketing category that could be used by an insurance company to pinpoint a potential customer as high risk.

Many researchers argue, however, that there are legitimate scientific uses for such data (see *Nature* **488**, 448–450; 2012). Jarmin says that the Census Bureau is exploring the use of data from credit-card companies to monitor economic activity. And researchers funded by the US National Science Foundation are studying how to use public Twitter posts to keep track of trends in phenomena such as unemployment.

But not everyone makes the distinction between commerce and academia, says Lane. "People conflate the concern about big data being used for private-sector purposes to make money with big data being used for research." In March 2014, for instance, while aiming to

"IT SHOULD BE HARD TO GET ACCESS TO DATA, BUT IT'S VERY IMPORTANT THAT SUCH ACCESS BE MADE POSSIBLE."

privacy sometimes complicates researchers' lives, it is necessary to uphold the public trust that makes the work possible.

"Difficulty in access is a feature, not a bug," she says. "It should be hard to get access to data, but it's very important that such access be made possible."

Many nations collect administrative data on a massive scale, but only a few, notably in northern Europe, have so far made it easy for researchers to use those data.

In Denmark, for instance, every newborn child is assigned a unique identification number that tracks his or her lifelong interactions with the country's free health-care system and almost every other government service. In 2002, researchers used data gathered through this identification system to retrospectively analyse the vaccination and health status of almost every child born in the country from 1991 to 1998 — 537,000 in all. At the time, it was the largest study ever to disprove² the now-debunked link between measles vaccination and autism.

Other countries have begun to catch up. In 2012, for instance, Britain launched the unified UK Data Service to facilitate research access to data from the country's census and other surveys. A year later, the service added a new Administrative Data Research Network, which has centres in England, Scotland, Northern Ireland and Wales to provide secure environments for researchers to access anonymized administrative data.

In the United States, the Census Bureau has been expanding its network of Research Data Centers, which currently includes 19 sites around the country at which researchers with the appropriate permissions can access confidential data from the bureau itself, as well as from other agencies. "We're trying to explore all the available ways that we can expand access to these rich data sets," says Ron Jarmin, the bureau's assistant director for research and methodology.

In January, a group of federal agencies,

significantly boost consumer privacy through a new data-protection regulation, the European Parliament proposed limiting the use of personal health data for research without specific consent, which would have severely curtailed researchers' access to those data. After objections from organizations such as the London-based biomedical-research charity the Wellcome Trust, the proposal looks likely to be jettisoned, but its fate will not become clear until 2016, when the final text of the regulation comes up for approval.

One solution to the privacy concerns has been to keep data under lock and key, tightly restricting who can access it. At the US research data centres, for instance, investigators are not allowed to take smartphones or flash drives into the rooms where they will use the centre's computer terminals. The computers themselves contain no data, but only link remotely to secure servers.

TECHNICAL ANSWERS

Computer scientists and cryptographers are experimenting with technological solutions. One, called differential privacy, adds a small amount of distortion to a data set, so that querying the data gives a roughly accurate result without revealing the identity of the individuals involved. The US Census Bureau uses this approach for its OnTheMap project, which tracks workers' daily commutes. Researchers at the bureau use actual data to build a statistical model based on where individual workers commute each day. They then build a synthetic data set that fits the model, but does not contain the actual data. This synthetic data set is released to the public, allowing users to draw accurate conclusions about transport and economic trends without tracking the exact movements of real individuals. Researchers are still learning to trust synthetic data, however, so few papers that have been published on this subject go beyond demonstrating the methods.

In any case, although synthetic data potentially solve the privacy problem, there are some research applications that cannot tolerate any noise in the data. A good example is the work showing the effect of neighbourhood on earning potential³, which was carried out by Raj Chetty, an economist at Harvard University in Cambridge, Massachusetts. Chetty needed to track specific individuals to show that the areas in which children live their early lives correlate with their ability to earn more or less than their parents. In subsequent studies⁵, Chetty and his colleagues showed that moving children from resource-poor to resource-rich neighbourhoods can boost their earnings in adulthood, proving a causal link.

Secure multiparty computation is a technique that attempts to address this issue by allowing multiple data holders to analyse parts of the total data set, without revealing the underlying data to each other. Only the results of the analyses are shared.

For instance, in 2010, the US Defense Advanced Research Projects Agency (DARPA)

“THE LESSON IS NOT TO UNDERESTIMATE PUBLIC CONCERNS. PUBLIC TRUST IS VERY FRAGILE.”

asked a team of cryptographers to develop a secure multiparty computation protocol to analyse the paths of commercial satellites and head off costly collisions. Currently, companies do this by sharing their orbit data, which they consider proprietary, to a trusted third party that performs the analysis. But DARPA concluded that secure multiparty computation could be used to predict possible collisions just as effectively, albeit a little more slowly.

In 2015, the Estonian company Cybertetica, based in Tallinn, said that it had used similar techniques to analyse financial filings of companies to detect tax fraud. It is also jointly analysing records from the country's tax and education ministries to explore whether university students who hold jobs fail their courses more often than those who focus exclusively on their studies.

There are still some problems in need of technical solutions — especially as government agencies look beyond their own walls. For instance, the Census Bureau wants to combine its internal data on the formation and activities of companies with public data on patents to examine the factors that drive corporate innovation. But it could be relatively easy to unmask the identities of companies involved in the analysis by matching them to information in the public patent database. Jarmin's team has not yet worked out an approach that adequately protects privacy.

But for the most part, technical solutions are now being put in place. Increasingly, what looks likely to hold up the research is a lack of clear ethical and legal guidance about how data on individuals can be used — for all purposes, including research.

Pam Dixon, executive director of the World Privacy Forum in San Diego, California, points to programmes such as India's national identification-card system, launched in 2010. This effort provided more than 900 million people with biometric identity cards that were linked to photographs, fingerprints and iris scans. The cards were supposed to be voluntary, and were used to identify rightful

recipients of social benefits such as fuel and unemployment aid.

But the country did not create a legislative framework to govern the use of the cards. They were soon discussed as gateways for a variety of essential services, such as salary payment and marriage registrations. This violated the original spirit of the programme, critics contended,

because data from the cards was not supposed to be coerced from individuals. The Indian Supreme Court ruled such uses of the system illegal on 11 August, but the country's Parliament has still not enacted a governing framework.

Likewise, in 2013, the United Kingdom launched the care.data programme to link records from patients' visits to general practitioners with their records from other parts of the health-care system, but there was no clear guidance on how the project's data were to be used. After it was revealed that the database designed to distribute patient data had inappropriately released some information to private entities — such as actuaries, which aid insurers in setting insurance rates — care.data came under fire. On 2 September, the National Health Service (NHS) said that the government will conduct a review of the security of NHS data and develop new opt-out and consent provisions. The system is intended to be available to all patients by 2016.

In the meantime, says Nicola Perrin, head of policy at the Wellcome Trust, the fallout has created huge delays in existing research projects, including clinical trials and health evaluation, audit and service research. Researchers in charge of SABRE, a large cohort study examining how diabetes and heart disease affect people of different ethnicities, have not received patient updates since March 2014; as a result, they risk sending requests for information to families whose loved ones may have died. The episode serves, for Perrin, as a cautionary tale about how the power of data could backfire if social unease with its uses is not addressed as soon as possible. “The lesson is to not underestimate public concerns,” she says. “Public trust is very fragile — it's difficult to build and easy to break.” ■

Erika Check Hayden is a reporter for *Nature* in San Francisco, California.

1. Chetty, R. *et al.* *Q. J. Econ.* **126**, 1593–1660 (2011).
2. Madsen, K. M. *et al.* *N. Engl. J. Med.* **347**, 1477–1482 (2002).
3. Chetty, R., Hendren, N., Kline, P. & Saez, E. *Q. J. Econ.* **129**, 1553–1623 (2014).
4. Ramirez, E., Brill, J., Ohlhausen, M. K., Wright, J. D. & McSweeney, T. *Data Brokers: A Call for Transparency and Accountability* (Federal Trade Commission, 2014).
5. Chetty, R., Hendren, N. & Katz, L. F. NBER Working Paper No. 21156 (2015); available at <http://www.nber.org/papers/w21156>



The mitochondria mystery

In the 1990s, French scientists wanted to see what happened to a mouse brain when they messed with the creature's mitochondria, the structures that generate energy inside most complex cells. The team looked at two mouse strains, called H and N, that carry slightly different mitochondrial-DNA sequences.

It was clear that the H mice learned to navigate mazes faster than their N cousins, but when the team swapped the mitochondria — creating H mice with N mitochondria and N mice with H mitochondria — their performance changed. Mitochondria from N seemed to slow down the learning process for H mice. N mice, meanwhile, improved slightly with H mitochondria¹. And

the team, led by geneticist Pierre Roubertoux at INSERM, the French National Institute for Health and Medical Research in Marseilles, found other changes in behaviour, and in brain anatomy, too.

The results came as a surprise, because such differences between mitochondrial genomes were seen as neutral — having no biological effect. “The long-held view was that the genetic variation we find within the mitochondrial genome doesn’t affect function,” says Damian Dowling, an evolutionary biologist at Monash University in Melbourne, Australia.

That view has been changing. A growing body of evidence suggests that mitochondria do not just produce energy, but also influence a wide range of cellular processes, from cell death to immune responses, and

The ‘powerhouses’ of the cell may have more roles than expected. Could that generate problems for mitochondrial replacement therapies?

By Garry Hamilton

that variations in the organelle matter very much. Variants in mitochondrial DNA are now linked to many common human conditions, including neurodegenerative diseases, cancer and ageing.

The effects of these variants may come about through the organelle’s long-evolved partnership with the much-larger nuclear genome. Studies in a handful of organisms have shown that just as for H and N mice, swapping healthy mitochondria between closely related strains can cause a mismatch between the genomes and can change important traits. The evidence, say Dowling and others, should raise questions about the safety

of a procedure that will soon be used in humans.

In February, the UK government approved mitochondrial replacement therapy, a technique that would allow a woman with a mitochondrial disorder to give birth to healthy children by pairing her nuclear DNA with the healthy mitochondria from a donor’s egg. The approval came after a 3.5-year effort to review the safety and ethics of creating individuals with DNA from three people (what some refer to as three-parent babies). And although many scientists lauded the decision, some worry that it is premature. “They’re not looking at the bigger picture,” says Ted Morrow, an evolutionary biologist at the University of Sussex in Brighton, UK, who is arguing for more-rigorous safety testing. “The

ILLUSTRATION BY SÉBASTIEN THIBAUT

standards for a shampoo seem to be harsher.”

A common refrain in favour of the therapy is that the genetic contribution from mitochondria is very small. And against the 3 billion base pairs of DNA and 20,000 genes found in the human nucleus, the mitochondrial genome can seem pretty insignificant (see ‘A complicated relationship’). Inherited solely through a mother’s egg, it comprises fewer than 17,000 base pairs and just 37 genes. But one cell can have thousands of copies of the mitochondrial genome, compared with just two of the nuclear genome — one from mum and one from dad.

Mitochondrial DNA also accumulates mutations incredibly fast, at about ten times the rate of nuclear DNA — and geneticists can use the resulting variation as a sort of molecular clock. The clock has allowed scientists to create a human family tree that shows several broadly related mitochondrial genomes, known as haplogroups, emerging in Africa somewhere around 150,000 years ago, including two that gave rise to the thousands of smaller haplogroups now found around the world.

The standing view was that the genetic differences between mitochondria in these groups were little more than a reflection of past migrations. But during the 1980s, researchers began to challenge that assumption. “Mitochondria control a central component of metabolism,” says David Rand, an evolutionary biologist at Brown University in Providence, Rhode Island. “So it followed that this variation ought to be very interesting.”

One way to examine whether mitochondria in one population work differently from those in another is to swap them. Such experiments would be unethical in people and impractical in many other animals, so Rand turned to fruit flies. He cross-bred two fly strains with different mitochondria and then repeatedly back-crossed them until the mitochondria from one were neatly paired with the nucleus of the other.

He then put fruit flies with similar nuclear genomes but different mitochondria together in a cage, and found that flies with specific mitochondrial genomes would quickly come to dominate the population². Something in the mitochondria was giving them a survival advantage. Subsequent work by Rand, Dowling and others has shown that it is not just the mitochondrial genome, but rather its interaction with the nuclear one that seems to be affecting a range of traits, including lifespan, reproductive success, rate of development, ageing, growth, movement, morphology and behaviour.

The findings extended beyond inbred laboratory animals such as fruit flies and mice. Over the past two decades, Ron Burton at the Scripps Institution of Oceanography in La Jolla, California, has found that cross-breeding closely related populations of tiny crustaceans known as copepods from tide pools on the Pacific coast often leads to a massive fitness breakdown for the animals³. Two clues led Burton to suspect that the reason was a mismatch between nuclear and mitochondrial DNA. First, the populations had very different mitochondrial genomes. Second, energy production was at the heart of all the sickly organisms’ deficiencies.

The clincher came when Burton chose females from the unhealthy animals and mated them with males from the same population as the females’ mothers. The resulting offspring, which once again had a natural combination of mitochondrial and nuclear genomes, were healthy. “That’s pretty striking,” says Burton. “And we did it with multiple different crosses.”

Extending these results to mammals has been difficult: Roubertoux’s mitochondrially mismatched mouse lines took more than 20 generations and 12 years to develop. But there are a few studies that have found similar results. Douglas Wallace, who heads the Center for Mitochondrial and Epigenomic Medicine at the Children’s Hospital of Philadelphia, combined the nucleus from a lab-mouse strain with mitochondria from a mouse known to contain two different, but normal, mitochondrial

genomes. His group found that the modified mice had altered circadian rhythms — the natural oscillations that follow a roughly 24-hour cycle — performed worse in mazes and seemed more stressed in certain experimental conditions, compared with unmodified animals⁴.

In humans, there is only indirect evidence that the common variation found in the mitochondrial genomes of healthy individuals could have biological effects. Certain mitochondrial haplotypes have been linked to disorders such as type-2 diabetes, Parkinson’s disease and cancer, and normal variation in the mitochondria is thought to influence general physical traits such as longevity and elite athleticism⁵.

“Correlations are just correlations,” says Göran Arnqvist, an evolutionary biologist at Uppsala University in Sweden, “but there’s now a large enough number of them to in itself provide ample evidence that there’s something going on with mitochondrial DNA”.

Powerhouse pairing

The question remains exactly how these variations could affect such a broad range of biological functions. Part of the answer seems to lie in their ties with the nuclear genome. Roughly 1,500 nuclear genes are

involved in mitochondrial function, including around 76 that encode proteins which bind to mitochondrially derived peptides.

Common variants could alter how these proteins interact. If a mitochondrially derived protein needs to fit snugly against a nuclear counterpart, even tiny changes in one partner could disrupt that binding, a possibility supported by 3D modelling^{6,7}.

A study published in 2009 compared mitochondria from two common human European lineages, called haplogroups J and H, in cells with the same nuclear DNA⁸. It showed that cells with haplogroup J mitochondria contained more than twice as many copies of mitochondrial DNA as those with haplogroup H, a difference that would be expected to have a big influence on the production of

mitochondrial proteins.

Such effects could alter the rate at which mitochondria supply energy, with consequences for many cellular activities. But emerging evidence points to other ways that mitochondria could have broad biological implications.

Various molecules created during the production of energy, such as free radicals, may have a direct influence on processes involved in ageing, inflammation and in some basic cell functions. And in May, a team of researchers led by Gerald Shadel at Yale University in New Haven, Connecticut, showed in mice that mitochondrial DNA can itself trigger an innate immune response against viral infection⁹. “They’re not just power factories,” says Rand. “They’re also in a sense a nerve centre, a thermostat for the cell and how it’s doing.”

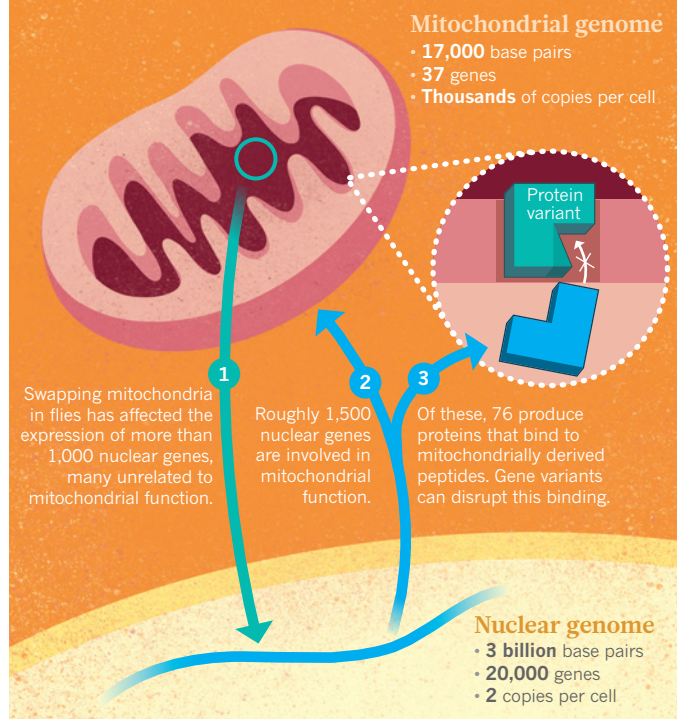
Researchers have also found evidence for a new class of mitochondrially derived peptide that might be encoded by sequences in other mitochondrial genes. One of these is humanin, a small peptide discovered by Japanese researchers in 2001 that increases sensitivity to insulin in diabetes-prone rats and mice¹⁰. The gene that encodes it is thought to reside in the mitochondrial gene for 16S ribosomal RNA. In March, researchers in the United States found a second potential example, MOTS-c, which is encoded by a small stretch of DNA tucked away in another gene. MOTS-c functions like a hormone, and when injected into mice helps to enhance insulin sensitivity and protect against obesity¹¹.

Some researchers now suspect that mitochondrial DNA produces a vast array of biologically active molecules — other small peptides as well as short stretches of RNA — that are part of a network of cross-communication between the mitochondrial and nuclear genomes. “The very viability of complex life — eukaryote life — depends on a really coordinated,

“They’re not just power factories, they’re also in a sense a nerve centre, a thermostat for the cell and how it is doing.”

A complicated relationship

The mitochondrial genome has evolved in concert with the nucleus of complex cells for hundreds of millions of years. Evidence suggests that even slight disruption of that relationship could have unexpected effects.



intimate set of interactions between these two genomes,” Dowling says. It is a partnership that has shaped and been shaped by aeons of evolution.

Given how well evolution has tuned this communication, many biologists are concerned about disrupting it in mitochondrial replacement therapy. The results of mitochondria-swapping experiments in other organisms, they say, should not be overlooked. “We haven’t seen anything fundamentally different between flies and humans in terms of interactions between the mitochondria and the nucleus,” says Klaus Reinhardt, an evolutionary biologist at the University of Tübingen in Germany.

The health effects may not be dramatic, says Burton, and they might not become apparent until decades after birth. “But I think there’s a definite possibility that you’d see things like disrupted fertility function, various forms of metabolic syndromes and changes in things that relate to metabolism in general.”

Call for caution

Reinhardt, Dowling and Morrow outlined their concerns in a 2013 paper¹² in *Science*. They called for studies aimed at addressing how mammals born after mitochondrial replacement fare in adulthood, and argued that scientists should at least look into haplotype matching — ensuring that the mitochondria from the donor and recipient come from the same haplogroup before transplant. Moving ahead at this juncture, they argued, “would place an experimental risk on families.”

But other researchers disagree. Scientists at Newcastle University, UK, and at Oregon Health & Science University (OHSU) in Beaverton, two institutions that pioneered mitochondrial replacement therapies, pointed to perfectly healthy macaque monkeys born at OHSU in 2009 after the procedure¹³.

They also pointed out that most of the evidence for risk stems from studies that used strains of flies and mice that had been highly inbred — a process that would increase the genetic differences between the strains and therefore produce a greater ‘mismatch’ when the mitochondria are swapped. They argued that such studies have little relevance for

human populations that interbreed all the time. The “lack of any reliable evidence of mitochondrial–nuclear interaction as a cause of disease in human outbred populations”, they wrote, “provides the necessary reassurance to proceed”.

Doug Turnbull, who heads the Newcastle group, also argues that correlations between different human mitochondrial haplotypes and common diseases are not definitive. “If we’re struggling to find a signal,” he says, “is that really something that’s likely to cause major difficulties?”

Ultimately, government approval hinged on a 2014 report prepared by a scientific review panel set up by the Human Fertilisation and Embryology Authority (HFEA), the body that regulates assisted-reproduction treatments in the United Kingdom. The panel’s chair, Andy Greenfield of the Medical Research Council, would not comment for this story, but the HFEA provided a written response to questions. It stated that deliberations were “time-consuming and as complex as the data themselves”, adding that most respondents presenting evidence to the panel viewed these issues as “at best minor or non-existent”. In its final report, the panel recommended that haplogroup matching be considered “as a precautionary step”. But it also stated that the benefits of doing so are “likely to be minimal”.

Some of the critics of the decision grant that mitochondrial replacement may be worth the risks for women who want to avoid passing rare and devastating disorders on to their children. Many, however, think that more time is needed to assess the risks. There is also concern that proponents of the therapy trivialized the role of mitochondria — particularly by likening mitochondrial replacement to changing the batteries in a camera. Critics argue that a failure to appreciate all the other processes in which the organelle is involved could lead to inadequate controls and wider application of mitochondrial replacement in fertility clinics.

“You may have a few thousand people who suffer from mitochondrial diseases,” says David Keefe, a reproductive biologist at New York University’s Langone Medical Center. “There are tens of millions of women who have infertility who may see this as a way to have the batteries charged in their eggs.”

At least one clinic in the United States has used cytoplasm from donor eggs to ‘normalize’ the eggs of women being treated for infertility, starting in the late 1990s (see *Nature* **509**, 414–417; 2014). The procedure, which probably transferred mitochondria as well, resulted in 17 births before the US Food and Drug Administration requested safety studies and the clinic stopped offering the procedure in 2001. Little is known about the health of the children born as a result of the procedure.

Turnbull rejects the slippery-slope argument. “In the UK, the legislation is very clear that mitochondrial donation can only be used to prevent serious mitochondrial disease,” he says. “I do not think there is any good evidence it would be useful for anything else.”

Although no one knows what the rapidly growing field of mitochondrial research will uncover next, both sides agree that there is no way to say for sure what will happen when doctors swap mitochondria in humans, short of actually doing it. For Dowling, at least, it is one scientific debate that he would rather not win. “I’d like to see this work so female sufferers of mitochondrial disease can have unaffected children,” he says. “So I hope we’re wrong.” ■ [SEE EDITORIAL P.425](#)

Garry Hamilton is a science writer in Seattle, Washington.

1. Roubertoux, P. L. *et al. Nature Genet.* **35**, 65–69 (2003).
2. Hutter, C. M. & Rand, D. M. *Genetics* **140**, 537–548 (1995).
3. Ellison, C. K. & Burton, R. S. *Evolution* **62**, 631–638 (2008).
4. Sharpley, M. S. *et al. Cell* **151**, 333–343 (2012).
5. Hudson, G., Gomez-Duran, A., Wilson, I. J. & Chinnery, P. F. *PLoS Genet.* **10**, e1004369 (2014).
6. Osada, N. & Akashi, H. *Mol. Biol. Evol.* **29**, 337–346 (2012).
7. da Fonseca, R. R., Johnson, W. E., O’Brien, S. J., Ramos, M. J. & Antunes, A. *BMC Genomics* **9**, 119 (2008).
8. Suissa, S. *et al. PLoS Genet.* **5**, e1000474 (2009).
9. West, A. P. *et al. Nature* **520**, 553–557 (2015).
10. Muzumdar, R. H. *et al. PLoS ONE* **4**, e6334 (2009).
11. Lee, C. *et al. Cell Metab.* **21**, 443–454 (2015).
12. Reinhardt, K., Dowling, D. K. & Morrow, E. H. *Science* **341**, 1345–1346 (2013).
13. Chinnery, P. F. *et al. PLoS Genet.* **10**, e1004315 (2014).

COMMENT

POLICY To tackle climate change, democracy must be enhanced not jettisoned **p.449**

EXHIBITION Spine-tingling show celebrates Soviet space programme **p.452**



THEATRE How is Nicole Kidman's turn as Rosalind Franklin? **p.454**

OUTREACH Popular uprising of science communication in Middle East **p.455**

REDOX POWER SYSTEMS



The Redox Cube is a planned 25-kilowatt fuel cell to be run on natural gas.

Reimagine fuel cells

Combine energy generation and storage to ensure that networks remain robust as more renewable technologies are adopted, urge **John P. Lemmon**.

Investments in solar photovoltaics and wind turbines are soaring as costs fall and governments and companies seek to reduce greenhouse-gas emissions. But fluctuating power from the wind and sun threatens to destabilize electricity grids. As more intermittent sources are connected, the power surges and crashes. This increases variability in voltage, in power and in the frequency of alternating current.

Already, Germany, which produces more

than 25% of its energy from renewables, is experiencing problems. Voltage glitches, by tripping crucial components and destroying equipment in factories and plants, have caused hundreds of thousands of euros of damage¹. Using coal plants to maintain stability adds greenhouse-gas emissions.

Generation and load must be balanced. Three approaches are in operation: using real-time demand and pricing incentives to control load; ramping natural-gas plants up

or down to compensate for fluctuating power; and storing energy. Each has downsides. Repeated requests to reduce demand agitate users and may be manipulated by third parties who stand to profit. Gas turbines and batteries cannot provide rapid (less than a second) high-power responses and supply energy for long periods. Batteries degrade and are expensive to replace. Combinations of batteries require multiple sets of electronics and control systems.

New types of fuel cell on the horizon could eliminate the need for such trade-offs and ease the integration of renewables into the grid. Currently, fuel cells are used to generate only electricity and heat. They can be modified to store energy and produce liquid fuels such as methanol, thanks to breakthroughs in materials and designs. Developing fuel cells with a battery mode is one focus of the programme I direct at the US Advanced Research Projects Agency–Energy (ARPA-E). I lead 13 projects across academia, industry and national laboratories.

Researchers must now demonstrate that fuel cells can perform multiple functions and still generate power efficiently.

INTEGRATION CHALLENGE

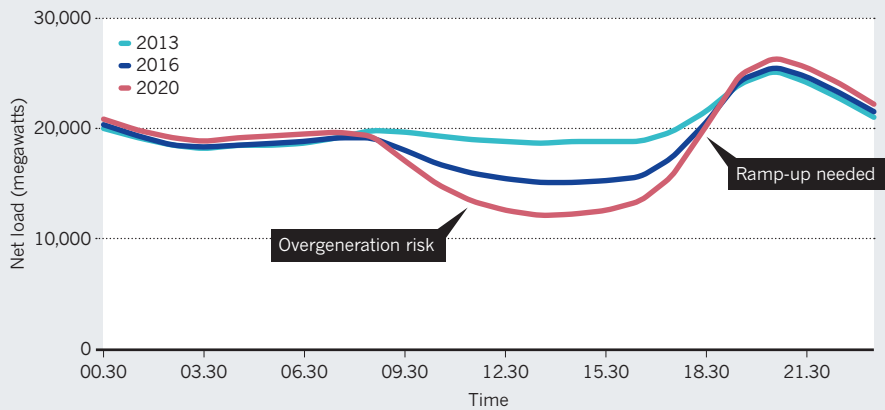
Power generation from distributed sources is expanding rapidly. Up to three times the current capacity of solar photovoltaics is projected² to come online in the United States by 2040 (an extra 15–75 gigawatts; GW). But conventional power grids are not designed to handle hundreds of thousands of small variable sources.

Renewables have two types of intermittency. First, output may fluctuate randomly moment-to-moment — for example, because of clouds casting shadows on solar panels. Second, output varies predictably through the day and night. According to one projection by the California Independent System Operator, the power needs for a late California afternoon in March 2020, when the sun is setting and people are returning home, will require a 13 GW ramp up over 3 hours — that is equivalent to switching on more than 20 power plants of 600-megawatt capacity (see 'Daily load').

Actions to decrease demand — such as changing thermostat settings by a few degrees — are often the first deployed when the grid is pushed to its limit on, say, a ▶

DAILY LOAD

The difference between power demand and generation varies throughout the day as sunshine and wind change and people travel to and from home. More power plants must be switched on in the evening. This imbalance is projected to increase as more renewables are used.



► sweltering summer day in a large city. Such ‘demand response’ can potentially shed up to 9% of the peak load in the United States³. But policy uncertainties make demand response fraught for grid system operators.

In May 2014, the US Court of Appeals struck down a 2011 order by the Federal Energy Regulatory Commission (FERC) that required grid operators to pay demand-response resources the full market price for energy, just as they would for conventional power generators. The court ruled that because demand response constitutes a retail market, it should be subjected to state (rather than FERC) jurisdiction. President Barack Obama appealed the ruling and the US Supreme Court will hear the case later this year. Either way, demand-side measures alone cannot accommodate the amount of renewables projected to be installed in the United States in the coming decades.

The latest natural-gas turbines are designed with a buffering capacity, to help to smooth out the power output from renewables. They get up to speed quickly, with ramping rates of tens of megawatts per minute. Still, it takes minutes — not seconds — to produce power from a standing start. Turbines that are already spinning can be brought into operation within seconds but with lower efficiency, higher emissions and, in the United States, 13–24% higher operating costs⁴.

The use of batteries for grid energy storage is receiving more attention. Energy companies are becoming increasingly confident that such devices are reliable in the field. Regulations are becoming more favourable. California has mandated 1.32 GW of storage capacity by 2020 and its utility companies have begun to buy storage devices.

One milestone was an announcement in May from the electric-car maker Tesla Motors, of Palo Alto, California. It has developed 7- and 10-kilowatt-hour (kWh) residential battery systems costing up to US\$3,500.

The company hopes that homeowners will want to avoid power cuts and paying peak electricity rates by charging the battery with solar panels or when grid prices are off-peak. But a battery’s storage capacity is limited by its size. Tesla’s 10 kWh, 1-metre-square, lithium-ion battery panel powers a home for one day, not enough for power cuts such as those following hurricanes and storms. Further improvements to cost and performance will come from higher-volume manufacturing of lithium-ion batteries and competing technologies such as redox (reduction–oxidation) flow batteries. No single technology provides the optimal balance.

Existing technologies such as improved wind-forecasting models, expansion of transmission infrastructure and other measures may help grids to remain stable as renewables penetration increases⁵. But larger grids (tens of megawatts or more) operating with more than 50% distributed generation are unprecedented. The adequacy of current technologies to support such a grid is unknown.

HYBRID FUEL CELLS

I back a different approach: fuel cells with built-in charge storage. Fuel cells have been touted for energy production for decades owing to their high electrical efficiencies. They have not been widely adopted because they are more expensive than combustion generators (\$3,000 per kilowatt, compared with \$1,000 per kilowatt). Over the past 15 years, US government-funded research programmes have aimed to lower those costs. Enhancing fuel-cell functionality would make them even more valuable.

Fuel cells are electrochemical devices similar to batteries that rely on a substance such as hydrogen or methane to produce power and heat. Two main types are in use: polymer fuel cells that operate at around 80 °C for vehicles; and solid-oxide fuel cells that operate above 650 °C for stationary

power. Polymer fuel cells require expensive platinum catalysts; solid-oxide fuel cells need expensive seals and connectors and have a limited lifetime. Adding functionality is difficult at either extreme of temperature.

Existing fuel cells are slow to respond to changes in current and voltage, taking more than a second. Hot fuel cells (operating above 650 °C) must avoid stressing their constituent materials, and their fuel processors take up to a minute to alter the rate at which methane is converted to hydrogen, carbon monoxide and carbon dioxide. Cells degrade if they are ‘starved’ of fuel.

Storing charge in or near a fuel cell’s electrode would speed up the device’s responsiveness, allow it to be ‘recharged’ without stopping and to live longer than conventional fuel cells. The battery-mode concept has been demonstrated in the lab in (cold) alkaline fuel cells with metal hydride anodes^{6,7} and a (hot) solid-oxide fuel cell with vanadium oxide electrodes⁸. Although the power densities reported were low, the cells stored voltage for minutes or hours. More materials research is required to increase the power output and reduce energy losses at the anode, which reacts with the fuel and stores charge.

INCREASING FUNCTIONALITY

Another function that could be integrated into fuel cells is the electrochemical conversion of natural gas (methane) into liquid fuels such as methanol. Current gas-to-liquid (GTL) technologies are economical only at large scales; Shell’s Pearl GTL plant in Qatar processes up to 45 million cubic metres of gas a day.

Rather than switching off solar panels or wind turbines when they are not needed, the spare electrons generated when supply exceeds demand could be directed towards

“Fuel cells have been touted for energy production for decades.”

making liquids for transportation fuels or chemicals. At natural-gas wells, such fuel cells could convert gas that would otherwise be flared or vented.

Such devices have been demonstrated in the lab in the past five years but are not yet commercially viable⁹. Researchers need to overcome the inherent stability of the methane molecule and convert it such that it is neither oxidized fully to CO₂ nor converted back to methane from an intermediate state.

I believe that both charge storage and fuel production can be incorporated cheaply into fuel cells that operate at intermediate temperatures. In the 200–500 °C range, many compounds capture and release hydrogen (magnesium hydride is one). These conditions are also more suitable for GTL conversion: hot enough for reasonable reaction kinetics, yet not so hot that methane oxidizes

completely to CO₂. Materials research on proton and oxygen ion conductors in the past decade shows that such fuel cells are possible.

For intermediate-temperature hybrid fuel cells to become a reality, researchers need to create solid electrolytes with high conductivity, and find electrode materials that have high activity and stability and that react with methane without forming coke (solid carbon). These devices must use less platinum catalyst and more impure fuel than low-temperature polymer cells; make do with cheaper seals and connectors; and last longer than higher-temperature solid-oxide fuel cells.

US researchers have made a start. I launched the Reliable Electricity Based on Electrochemical Systems (REBELS) programme at ARPA-E with \$33 million in funding in June 2014. It is starting to bear fruit¹⁰. Efforts elsewhere, particularly in Europe and Japan, are addressing hydrogen generation and GTL separately but could also benefit from hybrid fuel cells.

Researchers should prove the viability of intermediate-temperature fuel cells with these extra functions by demonstrating high power density and a lifetime of ten years, compared with current cell lifetimes of less than five. Cost savings must be validated through rigorous techno-economic modelling. Advances will then need to be scaled up from individual cells to kilowatt-scale systems, which will take 5–10 years.

Regulators, utility companies, technologists and users must define an appropriate mix of technologies and incentives to maintain the stability of the electricity grid in the coming decades. Hybrid fuel cells must be part of that conversation. ■

John P. Lemmon is a programme director at the US Department of Energy Advanced Research Projects Agency–Energy (ARPA-E), Washington DC, USA.
e-mail: john.lemmon@hq.doe.gov

1. Schröder, C. *Der Spiegel* 'Energy Revolution Hiccups: Grid Instability Has Industry Scrambling for Solutions' (16 August 2012).
2. US Energy Information Administration. *Modeling Distributed Generation in the Building Sectors* (EIA, 2013).
3. Federal Energy Regulatory Commission. *Assessment of Demand Response and Advanced Metering* (2012).
4. Lew, D. et al. *The Western Wind and Solar Integration Study Phase 2* (US Natl Renewable Energy Lab., 2013).
5. Weiss, J. & Tsuchida, B. *Integrating Renewable Energy into the Electricity Grid* (Brattle Group, 2015).
6. Chartouni, D., Kuriyama, N., Kiyobayashi, T. & Chen, J. *Int. J. Hydrogen Energy* **27**, 945–952 (2002).
7. Wang, C., Appleby, A. J., Cocke, D. L. *J. Electrochem. Soc.* **151**, A260–A264 (2004).
8. Van Overmeere, Q., Kerman, K. & Ramanathan, S. *Nano Lett.* **12**, 3756–3760 (2012).
9. Jacobs, T. J. *Pet. Technol.* **65**, 4135 (2013).
10. Duan, C. et al. *Science* <http://dx.doi.org/10.1126/science.aab3987> (2015).



Democracy is not an inconvenience

Climate scientists are tiring of governance that does not lead to action. But democracy must not be weakened in the fight against global warming, warns **Nico Stehr**.

There are many threats to democracy in the modern era. Not least is the risk posed by the widespread public feeling that politicians are not listening. Such discontent can be seen in the political far right: the Tea Party movement in the United States, the UK Independence Party, the Pegida (Patriotic Europeans Against the Islamization of the West) demonstrators in Germany, and the National Front in France.

More surprisingly, a similar impatience with the political elite is now also present in the scientific community. Researchers are increasingly concerned that no one is listening to their diagnosis of the dangers

of human-induced climate change and its long-lasting consequences, despite the robust scientific consensus. As governments continue to fail to take appropriate political action, democracy begins to look to some like an inconvenient form of governance. There is a tendency to want to take decisions out of the hands of politicians and the public, and, given the 'exceptional circumstances', put the decisions into the hands of scientists themselves.

This scientific disenchantment with democracy has slipped under the radar of many social scientists and commentators. Attention is urgently needed: ►

► the solution to the intractable ‘wicked problem’ of global warming is to enhance democracy, not jettison it.

VOICES OF DISCONTENT

Democratic nations seem to have failed us in the climate arena so far. The past decade’s climate summits in Copenhagen, Cancun, Durban and Warsaw were political washouts. Expectations for the next meeting in Paris this December are low.

Academics increasingly point to democracy as a reason for failure. NASA climate researcher James Hansen was quoted in 2009 in *The Guardian* as saying: “the democratic process doesn’t quite seem to be working”¹. In a special issue of the journal *Environmental Politics* in 2010, political scientist Mark Beeson argued² that forms of ‘good’ authoritarianism “may become not only justifiable, but essential for the survival of humanity in anything approaching a civilised form”. The title of an opinion piece published earlier this year in *The Conversation*, an online magazine funded by universities, sums up the issue: ‘Hidden crisis of liberal democracy creates climate change paralysis’ (see go.nature.com/pqgyr).

The depiction of contemporary democracies as ill-equipped to deal with climate change comes from a range of considerations. These include a deep-seated pessimism about the psychological make-up of humans; the disinclination of people to mobilize on issues that seem far removed; and the presumed lack of intellectual competence of people to grasp complex issues. On top of these there is the presumed scientific illiteracy of most politicians and the electorate; the inability of governments locked into short-term voting cycles to address long-term problems; the influence of vested interests on political agendas; the addiction to fossil fuels; and the feeling among the climate-science community that its message falls on the deaf ears of politicians.

Such views can be heard from the highest ranks of climate science. Hans Joachim Schellnhuber, founding director of the Potsdam Institute for Climate Impact Research and chair of the German Advisory Council on Global Change, said of the inaction in a 2011 interview with German newspaper *Der Spiegel*: “comfort and ignorance are the biggest flaws of human character. This is a potentially deadly mix”.

What, then, is the alternative? The solution hinted at by many people leans towards a technocracy, in which decisions are made by those with technical knowledge. This can be seen in a shift in the statements of some co-authors of Intergovernmental Panel on Climate Change reports, who are moving away from a purely advisory role towards

policy prescription (see, for example, ref. 3).

We must be careful what we wish for. Nations that have followed the path of ‘authoritarian modernization’, such as China and Russia, cannot claim to have a record of environmental accomplishments. In the past two or three years, China’s system has made it a global leader in renewables (it accounts for more than one-quarter of the planet’s investment in such energies⁴). Despite this, it is struggling to meet ambitious environmental targets and will continue to lead the world for some time in greenhouse-gas emissions. As Chinese citizens become wealthier and more educated, they will surely push for more democratic inclusion in environmental policymaking.

Broad-based support for environmental concerns and subsequent regulations came about in open democratic argument on the value of nature for humanity. Democracies learn from mistakes; autocracies lack flexibility and adaptability⁵. Democratic nations have forged the most effective international agreements, such as the Montreal Protocol against ozone-depleting substances.

GLOBAL STAGE

Impatient scientists often privilege hegemonic players such as world powers, states, transnational organizations, and multinational corporations. They tend to prefer sweeping policies of global mitigation over messier approaches of local adaptation; for them, global knowledge triumphs over local know-how. But societal trends are going in the opposite direction. The ability of large institutions to impose their will on citizens is declining. People are mobilizing around local concerns and efforts⁶.

The pessimistic assessment of the ability of democratic governance to cope with and control exceptional circumstances is linked to an optimistic assessment of the potential of large-scale social and economic planning. The uncertainties of social, political and economic events are treated as minor obstacles that can be overcome easily by implementing policies that experts prescribe. But humanity’s capacity to plan ahead effectively is limited. The centralized social and economic planning concept, widely discussed decades ago, has rightly fallen into disrepute⁷.

The argument for an authoritarian political approach concentrates on a single effect that governance ought to achieve: a reduction of greenhouse-gas emissions. By focusing on that goal, rather than on the economic and social conditions that go hand-in-hand

with it, climate policies are reduced to scientific or technical issues. But these are not the sole considerations. Environmental concerns are tightly entangled with other political, economic and cultural issues that both broaden the questions at hand and open up different ways of approaching it. Scientific knowledge is neither immediately performative nor persuasive.

ENHANCE ENGAGEMENT

There is but one political system that is able to rationally and legitimately cope with the divergent political interests affected by climate change and that is democracy. Only a democratic system can sensitively attend to the conflicts within and among nations and communities, decide between different policies, and generally advance the aspirations of different segments of the population. The ultimate and urgent challenge is that of enhancing democracy, for example by reducing social inequality⁸.

If not, the threat to civilization will be much more than just changes to our physical environment. The erosion of democracy is an unnecessary suppression of social complexity and rights.

The philosopher Friedrich Hayek, who led the debate against social and economic planning in the mid-twentieth century⁹, noted a paradox that applies today. As science advances, it tends to strengthen the idea that we should “aim at more deliberate and comprehensive control of all human activities”. Hayek pessimistically added: “It is for this reason that those intoxicated by the advance of knowledge so often become the enemies of freedom”¹⁰. We should heed his warning. It is dangerous to blindly believe that science and scientists alone can tell us what to do. ■

Nico Stehr is a sociologist and founding director of the European Center for Sustainability Research at Zeppelin University in Friedrichshafen, Germany. e-mail: nico.stehr@t-online.de

1. Adam, D. *The Guardian* ‘Leading climate scientist: “democratic process isn’t working”’ (18 March 2009).
2. Beeson, M. *Environ. Politics* **19**, 276–294 (2010).
3. Hansen, J. et al. *PLoS ONE* <http://dx.doi.org/10.1371/journal.pone.0081648> (2013).
4. REN21. *Renewables 2015 Global Status Report* (REN21, 2015).
5. Runciman, D. *The Confidence Trap: A History of Democracy in Crisis from World War I to the Present* (Princeton Univ. Press, 2013).
6. Stehr, N. *Information, Power and Democracy, Liberty is a Daughter of Knowledge* (Cambridge Univ. Press, 2015).
7. Pierre, J. *Debating Governance: Authority, Steering, and Democracy* (Oxford Univ. Press, 2000).
8. Rosanvallon, P. *The Society of Equals* (Harvard Univ. Press, 2013).
9. Hayek, F. A. *Nature* **148**, 580–584 (1941).
10. Hayek, F. A. *The Constitution of Liberty* (Routledge, 1960).



A hypersonic cruise missile engine developed for the US Defense Advanced Research Projects Agency.

MILITARY SCIENCE

Inventions of war

Ann Finkbeiner assesses a study of DARPA, the agency that readies US technologies for coming conflicts.

The research arm of the US Department of Defense is called the Defense Advanced Research Projects Agency — a disconcerting combination of words. DARPA is small, but it predicts what future wars might look like and gets the necessary technologies ready. It has had stunning successes, including the foundations of the Internet, satellites for reconnaissance and global positioning (called Corona and Transit, respectively) and stealth technology. DARPA also has a particular character: created in reaction to the Soviet Union's surprise launch of the Sputnik satellite, it is meant to prevent “technological surprises” through high risk, high pay-off research. The story of this entity, in business for 57 years, should be all kinds of interesting.

But how to tell it? Much of DARPA's work is classified, so any history will necessarily be gappy. And whereas the agency's virtue

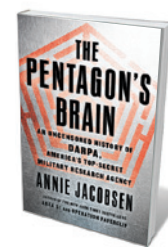
is its responsiveness to changes in warfare, that flexibility means that its history goes off in all directions. DARPA does not do research itself: it funds and works through a confusing variety of research centres, as well as industrial and academic bodies. The missions and titles of its internal offices vary with time. Even its name has switched periodically from ARPA to DARPA (‘Defense’ was first added in 1972, to emphasize the military nature of the agency's research). The result is a writer's nightmare, a story with a limitless cast of characters and no obvious storyline. In *The Pentagon's Brain*, journalist Annie Jacobsen has tackled the job by telling successive stories, focusing mostly on one programme at a time. Given DARPA's inherent complexity, this piecemeal approach leads to a struggle to capture the whole.

The bare bones of DARPA's story are as follows. It began in 1958, with two

physics-based cold war programmes: VELA, meant to detect nuclear explosions as verification for a treaty banning nuclear tests; and DEFENDER, which worked on defences and counter-defences against enemy missiles. By the time of the Vietnam War in the 1960s, war had become transformed into counter-insurgency, combating guerillas and converting citizens into collaborators. DARPA responded with AGILE, a programme that, in part, deployed psychology, sociology and anthropology to seek information on Vietnamese culture. When battlegrounds moved from jungles to cities, DARPA branched out into sensors and drones, then into computer-networked military operations, combating bioterrorism, robotics, artificial intelligence, human-machine interfaces and war simulations (C. Herzfeld *Nature* **451**, 403–404; 2008).

In Jacobsen's piece-by-piece telling, some of the stories are straightforward and firmly linked to DARPA. Others have gaps, probably resulting from information being classified, that she bridges with potentially related facts and speculation. In one of the straightforward tales, DARPA backed Project 137, a 1958 advisory meeting of academic scientists. One researcher, physicist Nick Christofilos, proposed an outlandish test to see whether nuclear explosions in the atmosphere could create a cloud of electrons that would be held by Earth's magnetic field for long enough to stop incoming missiles. (The test was done in 1958; the cloud lasted for weeks.) Another unambiguous example is SIMNET, a multi-player digital war game that DARPA created in the early 1980s to simulate realistic air and land battles. By mid-1990, SIMNET had become a rehearsal for a war in the Middle East, with desert terrain, cities, tanks, aircraft and armies; according to General Norman Schwartzkopf, it “eerily paralleled” the actual Gulf War, which started later that year.

For stories with holes, Jacobsen relies on implication, juxtaposing pieces of information that may or may not be related and trusting the reader to make the connections. Discussing DARPA's involvement in bioweap-



The Pentagon's Brain: An Uncensored History of DARPA, America's Top-Secret Military Research Agency
ANNIE JACOBSEN
Little, Brown: 2015.

onry, for instance, she assembles the following elements. In 1992, Ken Alibek, a highly placed refugee from Soviet bioweapons programme Biopreparat, briefed the Pentagon. DARPA recognized a need to invest in biology. In 1996, the agency kick-started a biowarfare programme. In 1997, it asked the group of academic science advisers known as JASON ►

► (A. Finkbeiner *Nature* 477, 397–399; 2011) to report on whether it was feasible to engineer pathogens to become more lethal (for example, by making aerosolized anthrax); an unclassified summary of JASON's report says it was. In 1999, Alibek became president of a company that aimed to find antidotes to bioweapons, which got a one-year contract from DARPA. I am unsure what I am meant to make of all this.

I am sure, however, that I am intended to view as ethically dubious DARPA's decade-old launch of programmes combining artificial intelligence, autonomous robots and brain–computer interfaces. Jacobsen cites a JASON report saying that any research in this area would be subject to ethical regulation. Then, referring to work published in *Nature* on the capacity of the hormone oxytocin to foster trust (M. Kosfeld *et al.* *Nature* 435, 673–676; 2005), Jacobsen wonders whether soldiers might be injected with the chemical to encourage them to trust robots. And after discussing DARPA's sponsored research into limb regeneration and perhaps even human cloning, Jacobsen speculates on whether DARPA is trying to create autonomous hunter-killer robots. Such argument-by-juxtaposition is effective in fiction. In non-fiction, it is unconvincing.

Ultimately, Jacobsen's focus on DARPA's programmes sidesteps the more intractable subject of what DARPA is. She never addresses such obvious questions as how DARPA stays ahead of the next war, and whether its flexibility and responsiveness have drawbacks, for example in the ratio of risk to pay-off. Furthermore, the book promises to “shine a light on DARPA's secret history” — secret because so many of the projects are classified. Yet the text, checked against sources, shows a certain amount of creative interpretation. I know from my own reporting on JASON that Jacobsen's chapter on the electronic fence in Vietnam inspired by the group has little to do with DARPA, and that her assessment of JASON as generally central to DARPA's programmes is exaggerated.

However flawed, *The Pentagon's Brain* is an exciting read that asks an important question: what is the risk of allowing lethal technologies to be developed in secret? Jacobsen worries that the technology that DARPA helps to create “may itself outstrip DARPA as it is unleashed into the world”. The prose might be caffeinated, but the message is serious, and has been since the first human picked up a rock and thought that it might be good for killing. ■

Ann Finkbeiner wrote *The Jasons and is co-proprietor of a science blog, The Last Word on Nothing.*
e-mail: anniekf@gmail.com



An ejection seat and suit used on the Soviet Vostok missions from 1961 to 1963.

SPACE TRAVEL

When Soviets ruled the great beyond

Tim Radford is thrilled by an unprecedented exhibition marking the USSR's cold war feats in space.

Between the cold war years of 1957 and 1966, the Soviet Union established primacy in space. Its heady list of triumphs embraces, in the 1950s alone, the first artificial object and first animal in orbit, and the first image of the far side of the Moon. In the next decade, it grew to include the first attempt on Venus, the first man in space, the first woman in space, the first three-man mission in space, and the first spacewalk, automaton touchdown on the Moon, lunar rover (1970), and scoop of Moon rock brought back to Earth by an automaton. Reflecting the significance and extent of those triumphs, the long-awaited *Cosmonauts at the Science Museum in London* assembles memorabilia and engineering marvels borrowed from around a score of Russian institutes.

► **NATURE.COM**
For more on science
in culture see:
[nature.com/
booksandarts](http://nature.com/booksandarts)

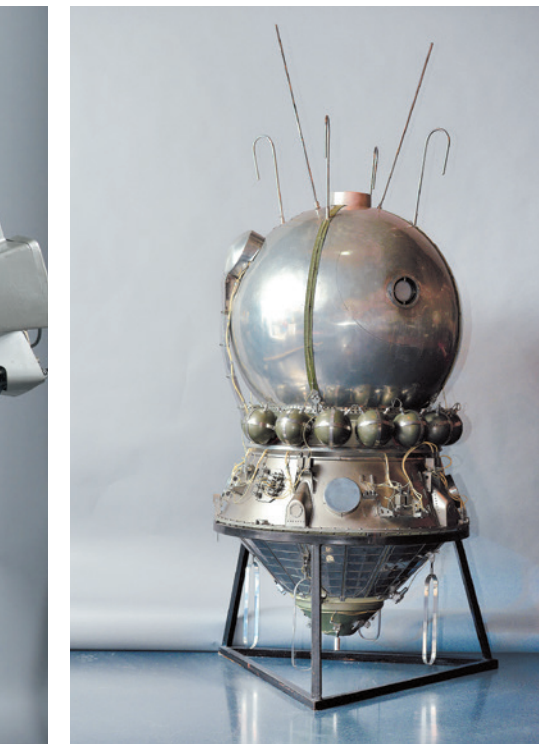
It opens with dreams: of high orbit, free fall and exit through an airlock, sketched on paper in 1933 by schoolmaster

Cosmonauts: Birth of the Space Age
Science Museum, London.
Until 16 March 2016.

and rocket visionary Konstantin Tsiolkovsky. It concludes with a recumbent mannequin in a cradle (a “tissue equivalent phantom” flown in 1969 to absorb and measure space radiation), representing the Soviet dream of a crewed mission to Mars, and a quotation attributed to Tsiolkovsky: “Earth is the cradle of humanity, but one cannot live in a cradle forever.” In between is a parade of hardware that none of us who followed the news greedily in those years had ever dreamed we might see assembled in one place, let alone in South Kensington.

The models are marvels. Here is a highly polished display model of Sputnik 1, launched in October 1957 (its chief designer, Sergei Korolev, reportedly said, “This ball will be exhibited in museums”). There are two engineering models: one of the two Lunokhod lunar rovers, the other of the once-secret lander *Lunniy Korabl*, designed to deliver

STATE MUSEUM AND EXHIBITION CENTER ROSIZO



A scale model of a 1960s Vostok spacecraft; a 1959 propaganda poster, *In the name of peace*; and a lower-body negative pressure suit from 1971.

one cosmonaut to the lunar surface in 1969. It flew, but not to the Moon, and the rest of us knew of its existence only two decades later.

And then there are the real things. Along with the charred, three-person *Voskhod 1* descent module used in 1964 is the descent module of *Vostok 6*. In it, cosmonaut Valentina Tereshkova orbited Earth for three days in 1963 before a return during which the heat shield was scorched by impact with Earth's atmosphere at 27,000 kilometres an hour. This is iconic stuff: the RD-108 engine that powered the space race; the complex space toilet designed to drain human waste aboard the space station Mir; the powered backpack with port and starboard lights for free flight beyond the spacecraft.

But what sets the scalp prickling are the little things that tell those other stories implicit in this dizzying show. There is Georgy Krutikov's 1928 drawing *Labour Commune*, a stratospheric dream prefiguring the great adventure. And there is a little metal mug once owned by Korolev, the man most people now recognize as the driver of the space race, and thus the hero of this story.

Korolev, a Ukrainian, had been incarcerated in a prison camp in the Kolyma region of Siberia during Joseph Stalin's notorious 1930s purges. No Westerner — and few Russians — knew his name during the cold war, so closed was the Soviet world. Fresh from wartime

labour detention, he arrived at the German Peenemünde base of the Nazi V-2 rocket programme to realize the dream of planetary exploration.

Sputnik 1 jolted Western complacency and helped to reignite the US space programme originally launched by the aerospace engineer and Nazi-turned-émigré Wernher von Braun. When Korolev died in 1966 during what should have been a routine operation, the new Soviet leader Leonid Brezhnev was a pallbearer. Even then, no one in the West knew of Korolev's existence.

Inevitably, the rocket engineer's genius surfaces again and again through the exhibition. There is a letter signed by Stalin authorizing the intercontinental ballistic-missile programme that made Sputnik 1 possible, and the personalized number plate YG1, used by Yuri Gagarin, the foundry worker who became a fighter pilot and, in 1961, the first man in space. There is Korolev's freehand drawing of the launch of canine cosmonauts Strelka and Belka. Alongside triumphant official Socialist realist posters there is a white lab coat daubed in red with the Russian for "Space is ours", a memento of a spontaneous 1961 celebration in Red Square. The pencils and sketch pad that Alexei Leonov took on his pioneering 1965

spacewalk — a near-catastrophe — are here, along with a later self-portrait of him floating at the end of a tether over the Black Sea.

The United States' role in the space race is hardly acknowledged, beyond a *Time* magazine cover declaring Soviet premier Nikita Khrushchev its 1957 man of the year. But the Soviet space effort seemed to lose momentum as the US Apollo programme — a story told in the Science Museum's main galleries — began in every sense to take off. Korolev's death must also have been a factor. The wonders went on, but the never-admitted race for the Moon was all but over.

This cosmic cornucopia reflects the intoxication of those first years and looks forward to the age of the space station. There is a spoon used aboard Mir by Sergei Krikalev, the man who went up as a Soviet cosmonaut and came down in 1992 as a citizen of the Russian Federation (and yes, there is a Soyuz descent module that carried a Mir crew back to Earth that year). But this unprecedented collection delivers more than a glimpse of distant exploratory technologies. It is a snapshot of Soviet history and, because the cold war warped the twentieth century, of global history, too. And where else could you see an ejector seat for a dog? The exhibits impose their own metaphors: see this show and be uplifted, transported, taken out of this world. It is the curatorial equivalent of a legal high. ■

The space suit worn by UK cosmonaut Helen Sharman on a 1991 mission to the Mir space station.

Tim Radford was science editor of *The Guardian* in London until 2005. e-mail: radford.tim@gmail.com

THEATRE

Lab's labour's lost

Philip Ball appraises Nicole Kidman's stage turn as crystallographer Rosalind Franklin.

The 1953 discovery of DNA's structure by James Watson and Francis Crick is a triumphant narrative with an uneasy subtext. Rosalind Franklin's crystallographic work was a vital part of the evidence. Yet, although her results (and those of Maurice Wilkins) were published in the same issue of *Nature* as theirs, Franklin was denied adequate credit for years (see *Nature* 496, 270; 2013). Watson and Crick never fully acknowledged the debt while she lived, and when she died at 37 of ovarian cancer, she effectively spared the Nobel committee the impossible decision of which trio to reward with the 1962 prize in medicine or physiology.

The question you need to ask yourself before seeing *Photograph 51*, Anna Ziegler's play about Franklin and the race to pin down the double helix, is how you like your science-in-theatre. Do you insist on adherence to the historical record, or do you accept that the aim is to illuminate and interrogate themes? There is plenty here to upset the stickler — not least, the status of the titular X-ray diffraction pattern of DNA, obtained by Franklin and PhD student Raymond Gosling at King's College London and used as evidence for Watson and Crick's double-helical model. Many of Ziegler's liberties (such as bringing Franklin's illness forward, and implying that Wilkins was infatuated with her) serve the narrative without compromising the core issues. But casting photograph 51 as a eureka moment is awkward.

Franklin did not fully interpret the image, for one thing. Nor did she take it (Gosling did), although that would not have been possible without her expertise. As Matthew Cobb writes in his excellent *Life's Greatest Secret* (Profile, 2015), the image's significance has often been overstated, largely because Watson chose to play it up ("The instant I saw the picture my mouth fell open") in his 1968 *The Double Helix* (Athenaeum).

In Watson's book there was also a hint, made much of in later accounts, of something underhand in how Wilkins — Gosling's supervisor — showed Watson the photo in early 1953. That was not true, although certainly Wilkins had clashed terribly with Franklin. Should Ziegler have used Watson's first-hand but unreliable narrative at face value to inform the plot? One might argue that if Watson could decorate the truth for the sake of a good story, why shouldn't she?

By adopting a dismissive tone towards Franklin, Watson's book inadvertently played



Nicole Kidman as Rosalind Franklin in *Photograph 51*.

a big part in launching her as a feminist icon. And Ziegler's play (which premiered in Los Angeles, California, in 2009) offers a more nuanced view of the myth.

Ziegler's players carry the story well. Watson (Will Attenborough) and Crick (Edward Bennett), naturally; the diffident Wilkins (Stephen Campbell Moore); Gosling (Joshua Silver) doing the PhD student's job of filling in gaps and making the tea, figuratively and literally. US structural biologist Donald Caspar (Patrick Kennedy) almost draws the work-obsessed Franklin into her first — and only — relationship. Linus Pauling, Max Perutz and Lawrence Bragg stay offstage. So does a fair bit of the science: we never see the double helix, and the audience is left to make what it will of phosphates being on the inside or the outside of the structure. That is no fault in itself — we are spared blackboard primers. But the metaphors about base-pairing (as Caspar takes Franklin's hand) or sexualized nestling of the twin strands are clunky.

The play belongs to Franklin. But she is written as so buttoned-up, prickly and focused that it is easier to warm to the urbane Crick or even the impetuous Watson. And casting a big star brings its own complications. Nicole Kidman's performance is restrained, but the glamour that attends her is the opposite of what the part demands. More surface ordinariness would have left room for a glimpse of depths.

Misogyny has loomed large in Franklin's

Photograph 51

ANNA ZIEGLER

Until 21 November 2015.

Noël Coward Theatre,
London.

tale ever since the feminist reading of Anne Sayre's *Rosalind Franklin and DNA* (Norton, 1975) — an

interpretation that Franklin would have disavowed, her sister has said. Had Franklin been less excluded and patronized by her male peers, might she have had the feedback and confidence to solve the structure first? In her authoritative *The Dark Lady of DNA* (HarperCollins, 2002), Brenda Maddox challenges that idea, suggesting that Franklin's class and religion (she came from a wealthy Jewish family) had an equal role in her isolation at King's. Ziegler finds a good accommodation: without any of the male characters becoming chauvinistic caricatures, we are left in no doubt that science was not welcoming to women in the 1950s.

More contentious in both history and the play is how to think about Franklin's science. Her experimental acumen is made clear; Kidman spends a lot of time at the lab bench. But what might have held Franklin back was that she did not trust model-building, believing that the structure must be revealed through mathematical analysis. Along with photograph 51, Watson and Crick assimilated other data, notably biochemist Erwin Chargaff's observation that in DNA, the amounts of adenine and thymine bases, and of cytosine and guanine, are equal. Perhaps more importantly, Watson, Crick and Pauling felt confident enough to foul up. All three committed howlers in trying to get the prize — Pauling's triple helix, published in early 1953, contained elementary errors. Ziegler's Franklin would have been mortified by such blunders.

That, perhaps, is the most valid message of *Photograph 51*. For science to thrive, there must be the freedom to fail. In Franklin's time, it is not surprising that a female scientist would think that she could ill afford that luxury. I am not at all sure that even a young Watson and Crick today could so freely take the risks they did. And shamefully, with evidence of gender imbalances in peer review and tenure, harassment and discrimination in the laboratory, and casual gender stereotyping still deemed acceptable by some leading scientists, the stakes remain still higher for a latter-day Franklin. ■

Philip Ball is a writer based in London.
e-mail: p.ball@btinternet.com

JOHAN PERSSON

Correspondence

Popular uprising spreads science

Hundreds of young Arab people are establishing initiatives to promote science in Arabic and raise scientific literacy across the Middle East, free of the censorship and bureaucracy of government and religious authorities (see *Nature Middle East* <http://doi.org/7p8>; 2015). They are publishing and translating scientific news and articles every day, including on topics such as evolution and sex education, which are widely taboo in many parts of the Middle East. Tens of thousands of reports, videos and infographics are popularizing a more objective way of thinking in the region. The movement relies on crowdsourcing from a vast pool of educated volunteers who are supervised by local scientists.

This 'uprising' has reached millions of people in a relatively short time. One science-communication group, Syrian Researchers, celebrated its millionth follower on Facebook in mid-2015 (www.syr-res.com). Another group, Scientific Saudi (www.scientificsaudi.com), has more than 250,000 social-media followers and is a learning partner of the Arabic edition of the *MIT Technology Review*.

Muath Alduhishy *University of Queensland, Herston, Australia.*
Mouhannad Malek Babrahman *Institute, Cambridge, UK.*
dr.alduhishy@gmail.com

Citizen projects can minimize conflicts

Well-structured schemes for citizen scientists can minimize the potential for conflicts of interest (*Nature* **524**, 265; 2015).

Projects such as the UK Breeding Bird Survey (go.nature.com/keyvpu), run by the British Trust for Ornithology, use volunteer-friendly protocols and specify sampling at representative locations to standardize volunteer commitment. The

primary motivator for observers is then whether to invest and participate in the survey, not whether they can influence which data are recorded.

Citizen scientists participating in well-structured schemes are more likely to deliver cost-effective monitoring on a large scale and to improve societal understanding of scientific issues. **James W. Pearce-Higgins** *British Trust for Ornithology, Thetford, UK.*

james.pearce-higgins@bto.org

Resolve ambiguities in China's emissions

As the former chair of the Consultative Group of Experts organized by the United Nations Framework Convention on Climate Change (UNFCCC) to help developing countries to produce carbon-emission inventories, I question the claim that China's emissions from coal have been overestimated (see *Nature* **524**, 276; 2015 and Z. Liu *et al.* *Nature* **524**, 335–338; 2015).

The accuracy of estimates depends largely on emission-factor estimates for the coal China uses (emission factor is the amount of carbon oxidized per unit of fuel consumed). For example, Liu and colleagues report emission factors that were estimated from the average carbon content of a range of high-quality to low-quality Chinese coal types. They write that these emission factors are 40% below the default values recommended by the 2006 guidelines of the Intergovernmental Panel on Climate Change (IPCC). However, I find their comparison flawed because the IPCC factor they use derives from coking coal, which contains more carbon and so has a higher emission factor than an 'average' coal type.

Furthermore, the conclusion by Liu *et al.* that China's fossil-fuel use in 2000–12 exceeded official figures by 10% seems incompatible with the authors' estimated emissions being 12%

lower than those calculated by the Chinese government. The official team's higher estimate was based on information from China's coal-quality database and from coal-trading contracts.

Such ambiguities call for clear resolution so that estimates of China's emissions are accurately conveyed.

Fei Teng *Tsinghua University, Beijing, China.*
tengfei@tsinghua.edu.cn

Overhaul rules for hazardous chemicals

The huge chemical explosion at the Chinese port of Tianjin on 12 August is another in the country's long list of industrial accidents involving chemicals. In 2010–14, more than 2,000 people were killed in 326 such accidents (J. Ren and Y. Mu *Chem. Enterp. Manag.* **16**, 28–31; 2015; in Chinese), calling into question the adequacy and enforcement of national regulations for the safe management of hazardous chemicals.

China needs better legislation and more detailed regulations for controlling risk at different stages in the life cycle of hazardous chemicals. Safety supervision must be made more effective, for example by drawing up regulations modelled on the European Union's Seveso Directive for industrial accidents (see go.nature.com/zyjp85). Supervision is currently disorderly, overseen by a fragmented and overlapping structure of multiple agencies, including the State Administration of Work Safety and the ministries of transport, public security, environmental protection, agriculture and health.

Effective enforcement of China's safety management of hazardous chemicals may be foiled by the limited expertise of company front-line managers (M. Liu *et al.* *Contemp. Chem. Ind.* **43**, 2661–2662; 2014; in Chinese). More data are

needed on the nature, handling and storage of the chemicals themselves so that risks can be properly assessed and managed (see go.nature.com/bzsydq).

Zhenwu Tang *North China Electric Power University, Beijing, China.*

Qifei Huang, Yufei Yang *Chinese Research Academy of Environmental Sciences, Beijing, China.*
huangqf@craes.org.cn

Olympics will make water scarcity worse

The 2022 Winter Olympics in Beijing threaten to seriously exacerbate water shortages in the area, where the available water per person is already only about 3% of the world's average (see also *Nature* **524**, 278–279; 2015).

The Winter Olympics will take place in February, when monthly precipitation in Beijing is less than 6 millimetres, so the games will need to rely exclusively on artificial snow (see go.nature.com/bkxbo8). This will entail pumping massive volumes of water out of reservoirs and rivers, further reducing the local population's water supply, and using huge amounts of energy to cool this water to make ice crystals that can be ejected. For example, producing artificial snow for winter sports in Beijing in 2010 used the annual equivalent of the water and electricity consumed by 8,300 and 5,400 households, respectively (see go.nature.com/ysdpbd; in Chinese).

Hong Yang *University of Oslo, Norway.*

Julian R. Thompson, Roger J. Flower *University College London, UK.*
hongyanghy@gmail.com

CONTRIBUTIONS

Correspondence may be sent to correspondence@nature.com after consulting go.nature.com/cmchno.

STAP cells are derived from ES cells

ARISING FROM H. Obokata *et al. Nature* **505**, 641–647 (2014) doi:10.1038/nature12968; retraction **511**, 112 (2014) doi:10.1038/nature13598; and H. Obokata *et al. Nature* **505**, 676–680 (2014) doi:10.1038/nature12969; retraction **511**, 112 (2014) doi:10.1038/nature13599

Two reports claiming a novel cellular reprogramming phenomenon, stimulus-triggered acquisition of pluripotency (STAP), were published in *Nature* last year^{1,2}, but then subsequently retracted^{3,4}. The identity of STAP cells and STAP-derived stem cells, however, has remained undetermined. Here we report the results of a whole-genome sequencing (WGS) investigation of STAP-related samples kept mainly at the RIKEN Center for Developmental Biology. We show that all purported STAP stem-cell lines were contaminated with embryonic stem (ES) cells, and that chimaeric mice and teratomas supposedly derived from STAP cells instead show ES cell contribution.

The original article¹ reported that exposure to low pH can reprogram differentiated cells into unique pluripotent cells (STAP cells),

from which two secondary cell lines were established; ES-like STAP stem cells and trophoblast stem-like Fgf4-induced stem cells capable of generating placental cells^{1,2}. Because STAP cells were not maintained as frozen stocks, we first performed WGS of 15 genomic DNA samples in total, including three representative STAP stem-cell lines with different genetic backgrounds, an Fgf4-induced stem-cell line, and seven ES cell lines established at the Wakayama laboratory before or during the STAP study (Extended Data Table 1). We determined genome-wide patterns of single-nucleotide polymorphisms (SNPs) that distinguish mouse strains 129/Sv (129) and C57BL/6 (B6), as well as green fluorescent protein (GFP) transgene types (Supplementary Methods and Extended Data Fig. 1a). No samples from the *Oct4*-GFP Fgf4-induced stem cells described in the original letter² were found (*Oct4* is also known as *Pou5f1*).

The STAP stem-cell line FLS and the Fgf4-induced stem-cell line CTS were reported to carry a homozygous insertion of a single *cag-gfp* transgene with the genetic background of 129 female × B6 male (Extended Data Table 1). However, these cell lines had co-insertions of two GFP transgenes⁵, sperm-specific acrosin-promoter-*gfp*⁶ and ubiquitously expressed *cag-gfp*⁷ (hereafter designated *Acr/cag-gfp*) at chromosome 3, which originated from an *Acr/cag-GFP* B6 mouse strain⁸ not described in the STAP papers^{1,2}. These STAP cell lines were then compared with four ES cell lines—FES1, FES2, and two nuclear transfer ES lines (ntESG1 and ntESG2) (ref. 9)—established from crossing the *Acr/cag-GFP* mouse strain with 129 mice in the Wakayama laboratory in 2005 (Extended Data Fig. 1a and Extended Data Table 1). FES1 and FES2 cells shared homologous SNP patterns with these STAP cell lines over the entire genome, including the 129 X chromosome, while ntESG1 and ntESG2 cells bearing B6 X chromosome were excluded from the comparison. Furthermore, these STAP cell lines shared two genomic characteristics with FES1, but not FES2; first, two chromosomal deletions (Fig. 1a) are present only in FES1 and all *Acr/cag-GFP* STAP stem-cell sublines, but not in the other cell lines examined, in the paternal *Acr/cag-GFP* mice (frozen stock in 2010), or in potential maternal 129 substrains available in Japan. Second, FES1 and the STAP cell lines with *Acr/cag-GFP* share

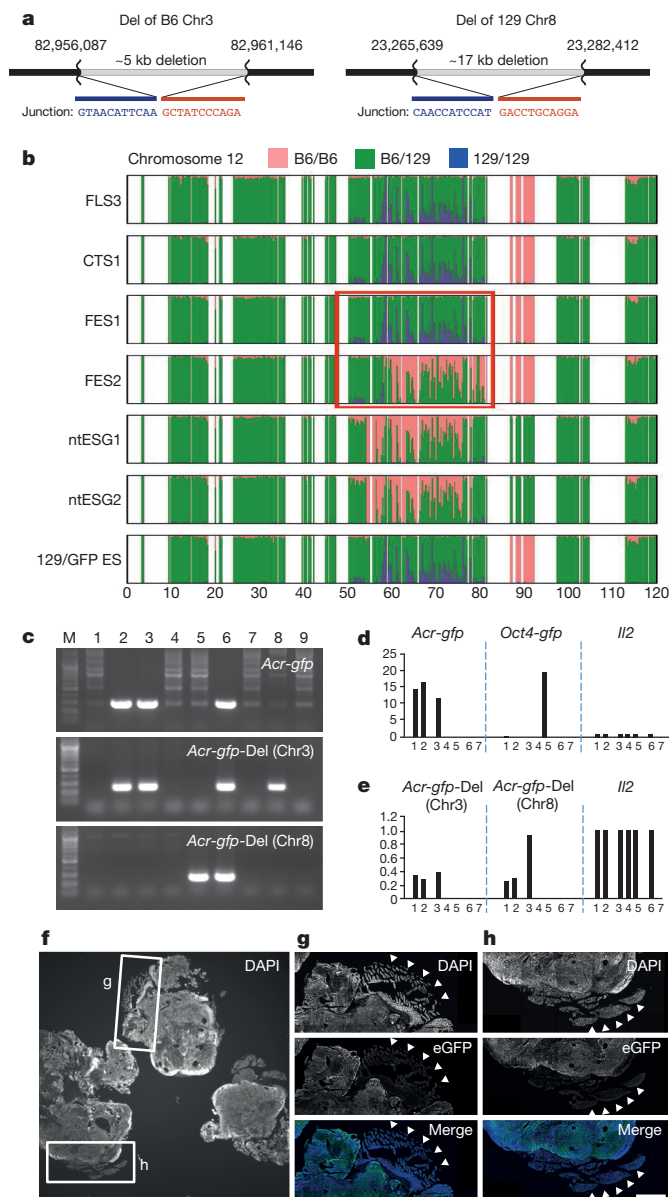


Figure 1 | STAP cells and STAP stem cells are derived from ES cells. **a**, Two genomic deletions exclusively shared by STAP and Fgf4-induced stem-cell lines and FES1 ES cells carrying *Acr/cag-gfp*. **b**, SNP mosaicism of chromosome 12 (Chr12) in *Acr/cag-GFP*⁺ cells. The panel shows SNPs in the 1-megabase resolution. A large SNP mosaic region (a red rectangle) is different between FES1 and FES2 ES cells. All *Acr/cag-GFP*⁺ STAP-cell lines have the same mosaicism as FES1 ES cells (see Extended Data Fig. 1 for Chr6 and Chr11). **c**, *Acr-gfp*, and two deletions (chromosomes 3 and 8) specific to FES1 ES cells are inherited into offspring from STAP cell chimaeric mice. **d**, **e**, Teratomas are derived from ES cells. qPCR reproducibly detects *Acr-gfp* (**d**), and FES1 ES-cell-specific deletions (**e**) in genomic DNAs prepared from the STAP cell teratoma paraffin block. Lanes 1: STAP cell teratoma; 2: STAP cell teratoma (separately prepared); 3: FLS4 (*Acr/cag-GFP*⁺ STAP stem cell); 4: 129B6F1 ES-5 (control ES cell); 5: GLS13 (*Oct4-GFP*⁺ STAP stem cell); 6: C57BL/6NCRSlc mouse; and 7: no template DNA. Each value shows fold-amplifications relative to the *Il2* gene (see Supplementary Methods). **f**, DAPI staining of a section taken from the STAP cell teratoma paraffin block. The intestinal epithelium and pancreatic tissue in the rectangles correspond Fig. 2e and Extended Data Fig. 4c from ref. 1, respectively. **g**, **h**, Magnifications of the rectangles with immunostaining for enhanced GFP (eGFP), indicating that these tissues are derived from GFP-negative host tissues (white arrowheads). Scale bar, 1 mm.

large SNP clusters that differ between FES1 and FES2 in three chromosomes (Fig. 1b and Extended Data Fig. 1b, c). These differential SNP clusters probably arose from chromosomal heterogeneity in the parental mouse colonies when FES1 and FES2 were established. It is highly unlikely that the *Acr/cag*-GFP STAP cell lines and FES1 all independently acquired these two unique deletions and inherited the same three mosaic chromosomes from parental mice. An ES cell stock, 129/GFP ES, was also found to share all these genomic features (Extended Data Table 1).

After the above three SNP clusters reflecting parental heterogeneity are excluded, the remaining 1,290 SNP alleles that distinguish FES1 and FES2 are supposed to have accumulated at or after establishment in 2005. Regarding these SNPs, STAP cell lines FLS3 and CTS1 and 129/GFP ES cells are nearly identical, but differ slightly from FES1 (at 30% of these alleles), suggesting that STAP cell lines FLS and CTS were derived from a sub-stock of FES1 ES cells.

The STAP stem-cell line GLS1–13 was reported as established from STAP cells prepared from genomic *Oct4* fragments (GOF) mice (B6 background) carrying the *Oct4-gfp* transgene¹⁰ in 2012. All these cell lines have a large truncation with a terminal inverted repeat in one of two X chromosomes (Extended Data Fig. 2a). An identical X chromosome was found in GOF-ES, an ntES cell line established from GOF mice in 2011, but not in parental GOF mice. It is unlikely that such a peculiar X chromosome abnormality would occur independently, strongly suggesting that the GLS lines were derived from the GOF-ES.

SNP analysis revealed that two independent STAP stem-cell lines, AC129-1 and AC129-2, had a 129B6F1 genetic background, while they were documented in the original article¹ as being established from 129 *cag*-GFP mice. We identified five heterozygous genomic anomalies: four deletions, and a duplication in these STAP stem-cell lines (Extended Data Fig. 2b, d), which were not found in the sequenced parental mouse genomes. We identified that these anomalies and sexual identity were shared by one of six control ES cell lines with *cag-gfp*, 129B6F1 ES1, established earlier than AC129. This is also the case for the other *cag*-GFP STAP stem-cell lines, FLS-T1 and T2, established in 2013. The 129B6F1 ES1 also shares a characteristic homozygous B6-SNP cluster in chromosome 6 with these four *cag*-GFP STAP stem-cell lines (Extended Data Fig. 2c, d). It is unlikely that the 129B6 ES1 line and these *cag*-GFP STAP stem-cell lines independently inherited all five chromosomal anomalies, the Y chromosome, and the same chromosome 6 from parental mice at establishment.

The article¹ describes 2N chimaeric mice generated from STAP cells bearing *cag-gfp* on the 129B6F1 background and their germ-line transmission (Fig. 4 and Extended Data Fig. 7 in ref. 1) as evidence for pluripotency. We found nine genomic DNA samples for the offspring of STAP cell chimaeric mice (Extended Data Fig. 7c in ref. 1). These contained not only the *Acr-gfp* insertion but also the two deletions unique to FES1-derived *Acr/cag*-GFP cell lines described above (Fig. 1c) indicating that the cells transmitted to the germ line in the chimaeric mice were derived from FES1 ES cells.

The article¹ also describes teratomas derived from *Oct4*-GFP STAP cells as evidence for pluripotency (Fig. 2 and Extended Data Fig. 4 in ref. 1). We found a glass slide specimen from which all these teratoma images were taken, and its corresponding paraffin block. Quantitative PCR of genomic DNA extracted from this paraffin block reproducibly indicated that these teratoma tissues formed from FES1-derived cells (Fig. 1d, e). Immunostaining revealed that intestinal epithelium tissue (Fig. 2e, right in ref. 1) and pancreatic tissue (Extended Data Fig. 4c in ref. 1), shown as teratomas from STAP cells¹, were GFP-negative and, thus, of host mouse origin (Fig. 1f–h).

Control genomic DNA sequences for STAP cell chromatin immunoprecipitation sequencing (ChIP-seq) experiments (Fig. 4 in ref. 2) had been deposited in the NCBI database². To gain sufficient sequencing coverage, we re-sequenced the genomic DNA prepared from the STAP cell lysate used for ChIP-seq (Extended Data Fig. 1a). We confirmed that this STAP cell sample shared all the genomic characteristics described above for 129B6F1 ES1 (Extended Data Fig. 2c), indicating that the STAP cell sample used for ChIP-seq was derived from 129B6F1 ES1 cells.

In summary, our investigations based on WGS of STAP-cell-related materials reveal that all of these materials are derived from previously established ES cell lines and refute the evidence shown in the two *Nature* papers^{1,2} that cellular stress can reprogram differentiated cells into pluripotent cells. Data described here were presented to an external investigative committee convened by RIKEN. Raw data are available at the DDBJ sequence read archive (DRA) under accession number DRA002862.

Daijiro Konno^{1*}, Takeya Kasukawa^{2*}, Kosuke Hashimoto^{2*}, Takehiko Itoh³, Taeko Suetsugu¹, Ikuo Miura⁴, Shigeharu Wakana⁴, Piero Carninci² & Fumio Matsuzaki¹

¹Cell Asymmetry, RIKEN Center for Developmental Biology, 2-2-3 Minatogima-Minamimachi, Chuo-ku, Kobe 650-0047, Japan.

email: fumio@cdb.riken.jp

²Division of Genomic Technologies, RIKEN Center for Life Science Technologies, RIKEN Yokohama Campus, Yokohama, Kanagawa 230-0045, Japan.

email: carninci@riken.jp

³Graduate School of Bioscience and Biotechnology, Tokyo Institute of Technology, 4259 Nagatsuta-cho, Midori-ku, Yokohama, Kanagawa, 226-8503, Japan.

⁴Japan Mouse Clinic, RIKEN BioResource Center, 3-1-1, Koyadai, Tsukuba-shi, Ibaraki 305-0074, Japan.

*These authors contributed equally to this work.

Received 23 January; accepted 20 July 2015.

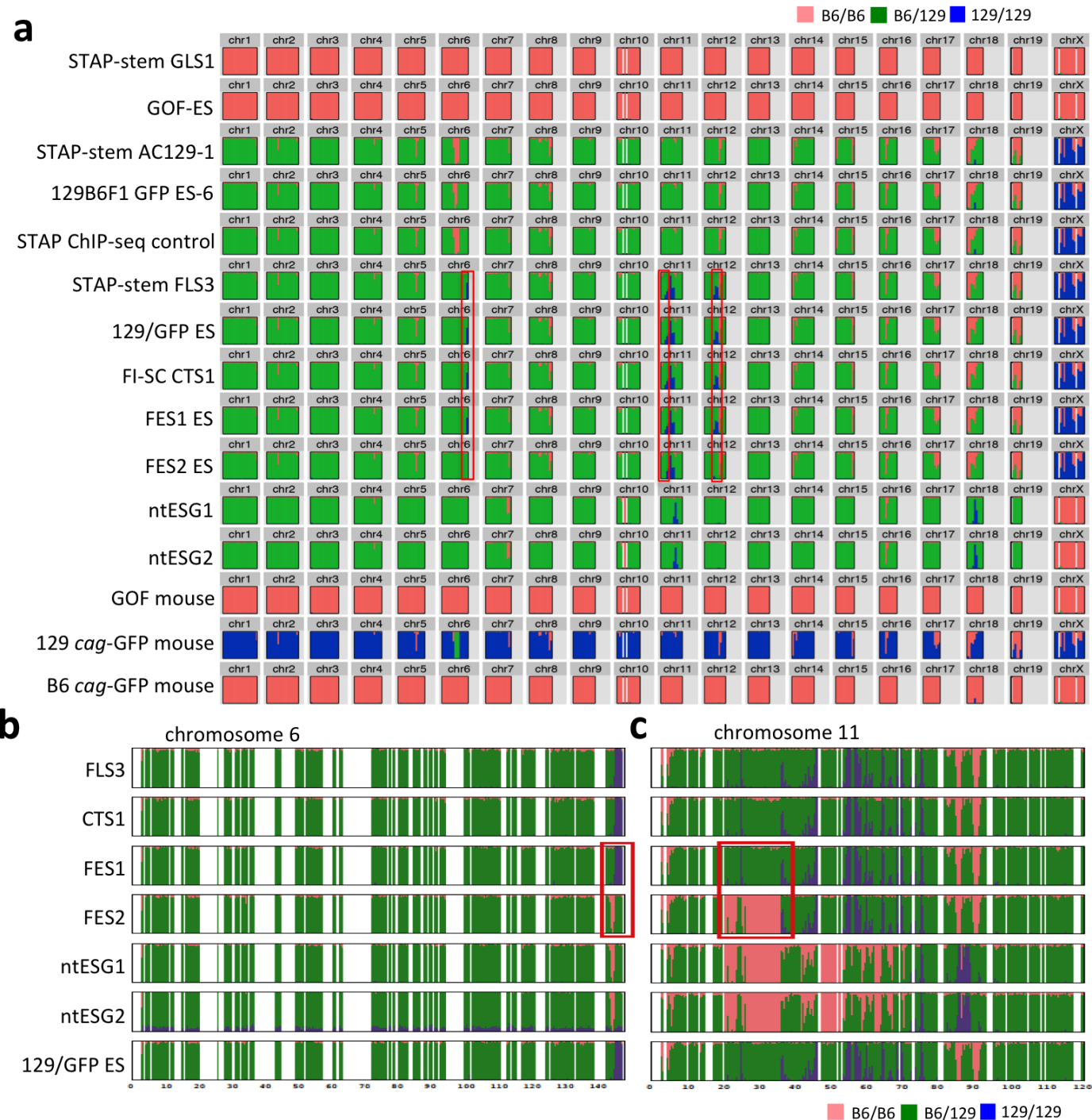
- Obokata, H. *et al.* Stimulus-triggered fate conversion of somatic cells into pluripotency. *Nature* **505**, 641–647 (2014).
- Obokata, H. *et al.* Bidirectional developmental potential in reprogrammed cells with acquired pluripotency. *Nature* **505**, 676–680 (2014).
- Obokata, H. *et al.* Retraction: stimulus-triggered fate conversion of somatic cells into pluripotency. *Nature* **511**, 112 (2014).
- Obokata, H. *et al.* Retraction: bidirectional developmental potential in reprogrammed cells with acquired pluripotency. *Nature* **511**, 112 (2014).
- Endo, T. A. Quality control method for RNA-seq using single nucleotide polymorphism allele frequency. *Genes Cells* **19**, 821–829 (2014).
- Nakanishi, T. *et al.* Real-time observation of acrosomal dispersal from mouse sperm using GFP as a marker protein. *FEBS Lett.* **449**, 277–283 (1999).
- Okabe, M., Ikawa, M., Kominami, K., Nakanishi, T. & Nishimune, Y. 'Green mice' as a source of ubiquitous green cells. *FEBS Lett.* **407**, 313–319 (1997).
- Nakanishi, T. *et al.* FISH analysis of 142 EGFP transgene integration sites into the mouse genome. *Genomics* **80**, 564–574 (2002).
- Ohta, H. & Wakayama, T. Generation of normal progeny by intracytoplasmic sperm injection following grafting of testicular tissue from cloned mice that died postnatally. *Biol. Reprod.* **73**, 390–395 (2005).
- Ohbo, K. *et al.* Identification and characterization of stem cells in prepubertal spermatogenesis in mice. *Dev. Biol.* **258**, 209–225 (2003).

Supplementary Information is available in the online version of the paper.

Author Contributions: D.K., T.K. and H.H. all contributed equally to this work. D.K. designed and performed PCR analyses of genomic DNA from cell and tissue samples. T.K. and K.H. designed WGS and analysed the data. T.I. performed detailed whole-genome SNP analysis. T.S. histochemically analysed teratoma samples. E.M. and S.W. performed TaqMan PCR analyses. P.C. designed and managed WGS and analysis. F.M. designed and organized this investigation, and wrote the manuscript.

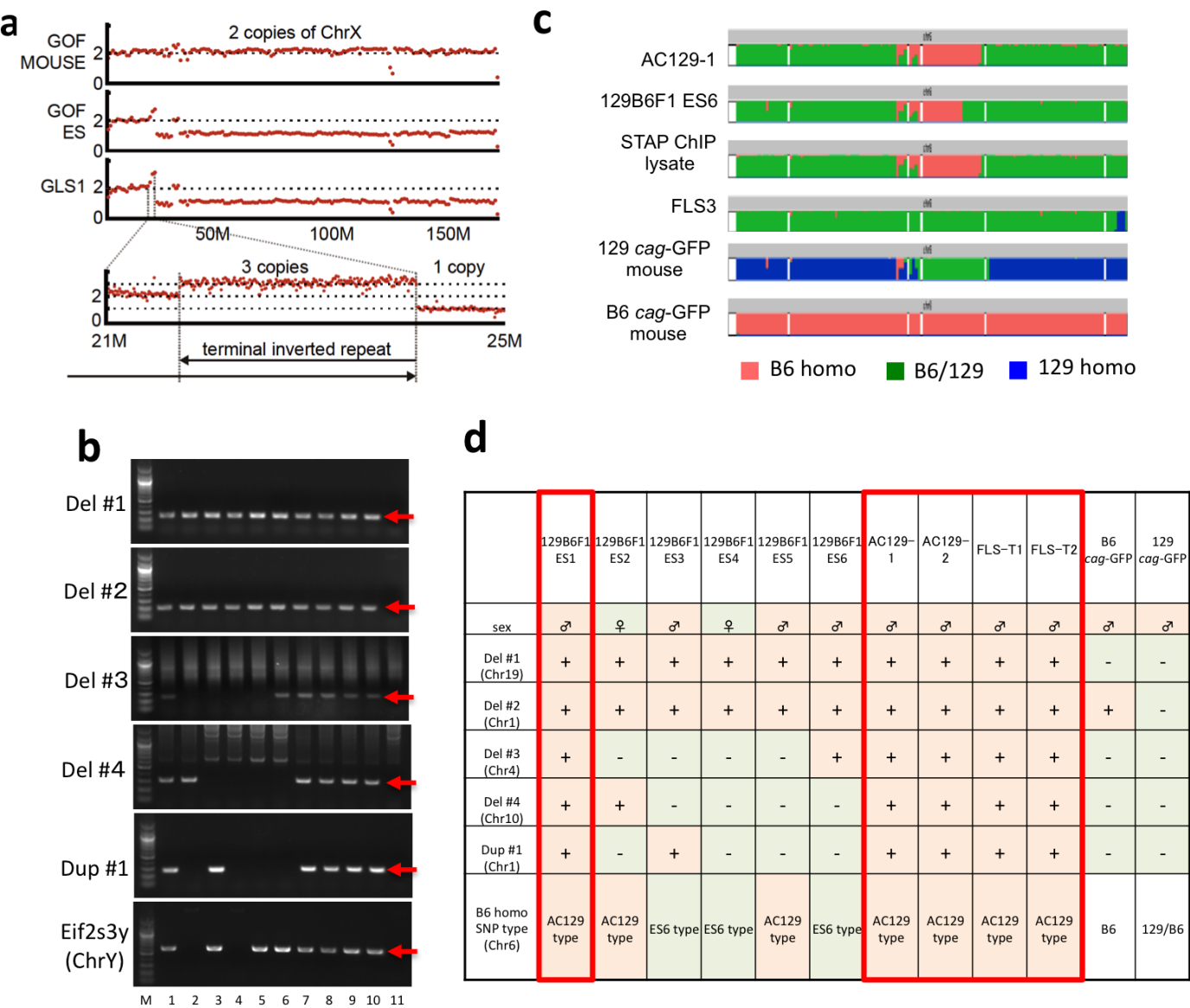
Competing Financial Interests Declared none.

doi:10.1038/nature15366



Extended Data Figure 1 | Genome-wide SNP patterns of STAP-cell-related cells and mice. **a**, SNP patterns of STAP stem cells, Fgf4-induced stem cells and related ES cells as revealed by WGS. Chromosomes 1–19 and X are aligned from left to right. All cell lines and mouse strains except for STAP stem cell GLS1, GOF-ES and GOF-mouse are male. 129/GFP ES cells and the re-sequenced control DNA of STAP cells for ChIP-seq (Fig. 4 in ref. 2) are also shown. B6-homozygous, B6/129-heterozygous and 129-homozygous SNPs are shown in magenta, green and blue, respectively. Note that ntESG1 and ntESG2 inherited the B6-type X chromosome from maternal mice. Genomic regions in which FES1 and FES2 ES cells have different SNP clusters in

chromosomes (chromosomes 6, 11 and 12) are marked by red rectangles. See **b**, **c** and Fig. 1b for a high-resolution map. SNP resolution is 10 Mb. **b**, **c**, High-resolution view of chromosomes 6 (**b**) and 11 (**c**), which show differential SNP clusters (red rectangles) between FES1 and FES2. In these regions, all SNPs are 129-type in FES1, and B6-type in FES2 on one of the homologous chromosomes. Therefore, in these regions, 129/B6 SNPs (green) and 129/129 SNPs (blue) in FES1 correspond to B6/B6 SNPs (red) and B6/129s (green) in FES2. STAP stem-cell line FLS3, Fgf4-induced stem-cell line CTS1, and 129/GFP ES cells share the same SNP patterns with FES1. SNP resolution is 1 Mb.



Extended Data Figure 2 | STAP stem cells derived from GOF-ES cells and *cag-GFP* ES cells. **a**, Copy number of chromosome X in STAP stem cells GLS1 and GOF-ES cells, both of which have *Oct4-gfp* transgenes with a B6 background. These lines have one very short X chromosome of ~23 Mb with a terminal inverted repeat and a normal X chromosome. **b**, PCR detection of chromosomal anomalies and Y chromosome in *cag-GFP* STAP stem-cell lines and parental mouse strains. Lanes 1–6: control ES cells, 129B6F1 ES1–6; 7: STAP stem cell AC129-1; 8: STAP stem cell AC129-2; 9: STAP stem cell FLS-T1; 10: STAP stem cell FLS-T2; and 11: GOF-ES. Deletions 1–4 and duplication 1 are located on Chr19: 32,857,093–32866,121, Chr1:140,698,249–140,702,693, Chr4:123,747,239–123,763,596, Chr10:43,265,147–43,267,270

and Chr1: 180,730,393–180,732,937, respectively. **c**, Distribution of B6-type and 129-type SNPs along chromosome 6. The B6-homozygous SNP cluster (magenta) in the middle, which probably arose from the inheritance of the parental 129, is heterogeneous in length among six control ES cell lines. The four *cag-GFP* STAP stem-cell lines share the same length of the B6 SNP cluster with control ES 129B6F1 ES1. Note that the 129/B6-heterozygous SNP region in the 129 *cag-GFP* mouse is longer than that of AC129-1. **d**, Table summarizing the chromosomal anomalies and differential types of Chr6 B6-homozygous SNP clusters in the *cag-GFP* cell lines and parental mice. Control ES cell 129B6F1 ES1 shares all the characteristic features with the four *cag-GFP* STAP stem-cell lines.

BRIEF COMMUNICATIONS ARISING

Extended Data Table 1 | STAP-cell-related cell lines

Cell line *	Cell type	GFP insertion †	Sex	Genetic background ‡	Established §	Origin	Used in
FLS1–8 (FLS3)	STAP stem cell	<i>Acr/cag-gfp</i> hetero (<i>cag-gfp</i> homo)	♂	129X1SLC ^Q /B6N SLC ♂	2012 1/31-2/2	FES1 ES	Article Fig. 5c, j-l, eFig. 8d,i,j
CTS-1, 11–13 (CTS1)	Fgf4-induced stem cell	<i>Acr/cag-gfp</i> hetero (<i>cag-gfp</i> homo)	♂	129X1SLC ^Q /B6N SLC ♂	2012 5/25, 7/9	FES1 ES	Letter Fig. 2f, g, i, Fig. 4, eFig. 2a, b, eFig. 3
GLS-1–13 (GLS-1)	STAP stem cell	<i>Oct4-gfp</i>	♀	B6	2012 1/31	GOF-ES	Article Fig. 5 a-c, eFig. 8k Letter Fig. 4
AC129-1, 2 (AC129-1)	STAP stem cell	<i>cag-gfp</i> homo	♂	129X1SLC ^Q /B6N SLC ♂ (129X1SLC) ¶	2012 8/13	129B6F1 ES1	Article Methods Letter Fig. 2i, 4
FLS-T1, T2 #	STAP stem cell	<i>cag-gfp</i> homo	♂	129X1SLC ^Q /B6N SLC ♂	2013 2/22	129B6F1 ES1	—
129B6F1 ES1–6 (129B6F1 ES6) ✕	fertilized egg ES cell	<i>cag-gfp</i> homo	♂♀	129X1SLC ^Q /B6N SLC	2012 4/19	—	Article Fig. 8d,i,j Letter Fig. 2, Fig. 4, eFig. 3,4,6
129/GFP ES **	ES cell	<i>Acr/cag-gfp</i> hetero	♂	129X1SLC ^Q /B6N SLC ♂	Unknown	—	—
GOF-ES	ntES cell	<i>Oct4-gfp</i>	♀	B6	2011 5/26-10/31	—	—
FES1 †	fertilized egg ES cell	<i>Acr/cag-gfp</i> hetero	♂	129X1SLC ^Q /B6N SLC ♂	2005 12/7	—	—
FES2 †	fertilized egg ES cell	<i>Acr/cag-gfp</i> hetero	♂	129X1SLC ^Q /B6N SLC ♂	2005 12/7	—	—
ntESG1 †	ntES cell	<i>Acr/cag-gfp</i> hetero	♂	B6N SLC ^Q /129 ^{Ter} CLEA ♂	2007 8/3	—	—
ntESG2 †	ntES cell	<i>Acr/cag-gfp</i> hetero	♂	B6N SLC ^Q /129 ^{Ter} CLEA ♂	2005 1/20	—	—

*The line subjected to WGS is indicated in parentheses in cases in which several sublines were established for one cell type. Other sublines were confirmed by PCR and sequencing.

†The type of *gfp* transgene described in the original *Nature* papers^{1,2} is shown in parentheses when it is different from the one determined by WGS.

‡The parental mouse strains and their genetic background heterogeneity were identified using the TaqMan PCR system that discriminated mouse substrains and breeders. Parental mouse genotypes are not homogeneous as judged by their SNPs. 129X1SLC denotes 129X1/SvJmsSlc; B6N SLC denotes C57BL/6NcrSlc; B6 denotes C57BL/6; 129+Ter CLEA denotes 129+Ter/SvJcl.

§Cultivation start date. For GOF-ES this is the date of cell line establishment, and for FES1, FES2, ntESG1 and ntESG2 this is the date when frozen stocks were created. Information was obtained from T. Wakayama and his laboratory members.

||Described as '*cag-gfp* (homozygous)' by the author who established the cell lines.

¶Described as '*129 cag-gfp* (homozygous)' by the author who prepared mice and established those cell lines.

#Obtained from T. Wakayama. WGS of FLS-T1 and FLS-T2 was not performed. PCR and Sanger sequence analyses showed that these two lines shared all characteristic genomic anomalies with AC129.

✕The genome of 129B6F1 ES6 was sequenced, but further analysis showed that 129B6F1 ES1 rather than 129B6F1 ES6 shares all genomic anomalies found in AC129-1 (see main text).

**129/GFP ES corresponds to *Acr/cag-GFP* cells found in the laboratory of the authors, and has identical genomic structures, including deletions and fine SNP patterns, to those of FLS1 and CTS1.

††Obtained from H. Ohta. Full names: FES1: 129B6GFP1 FES male; FES2: 129B6GFP2 FES male; ntESG1: 129B6F1G1; ntESG2: 129B6F1G2.

Failure to replicate the STAP cell phenomenon

ARISING FROM H. Obokata *et al.* *Nature* **505**, 641–647 (2014) doi:10.1038/nature12968; retraction **511**, 112 (2014) doi:10.1038/nature13598; and H. Obokata *et al.* *Nature* **505**, 676–680 (2014) doi:10.1038/nature12969; retraction **511**, 112 (2014) doi:10.1038/nature13599

Although the reports that stress (such as exposure to acid) can coax somatic cells into a novel state of pluripotency^{1,2} have been retracted^{3,4}, the validity of stimulus-triggered acquisition of pluripotency (STAP) remains unclear (<http://dx.doi.org/10.1038/protex.2014.008> and Supplementary Information). Here we describe the efforts of seven laboratories to replicate STAP, including experiments performed within the laboratory where STAP first originated, as well as re-analysis of the sequencing data from the STAP reports. Neonatal cells treated with two STAP protocols exhibited artefactual autofluorescence rather than bona fide reactivation of an *Oct4* (also known as *Pou5f1*) and green fluorescent protein (GFP) transgene reporter, did not reactivate pluripotency markers towards embryonic stem (ES)-cell-like levels, and failed to generate teratomas or chimaerize blastocysts. Re-analysis of the original RNA sequencing (RNA-seq) and chromatin immunoprecipitation sequencing (ChIP-seq) data identified discrepancies in the sex and genetic composition of parental donor cells and converted stem cells, and revealed a STAP-derived cell line to be a mixture containing trophoblast stem cells, attesting to the importance of validating the properties and provenance of pluripotent stem cells using a wide range of criteria.

To assess the reprogramming capacity of STAP protocols, we used a transgenic *Oct4*-GFP reporter, which shows GFP reactivation during *Oct4*/Sox2/Klf4 reprogramming, in established induced pluripotent stem (iPS) cells and in the gonads of mid-gestation ‘all iPS cell’ embryos generated by tetraploid complementation^{5–7} (Extended Data Figs 1 and 2a). Working within the Vacanti laboratory where the concept of STAP cells originated, and assisted by a co-author of the STAP papers, a Daley laboratory member (A.D.L.A.) attempted to replicate two reported STAP protocols: (1) mechanical trituration and acid treatment of mouse lung cells (Brigham and Women’s Hospital (BWH) protocol; see Supplementary Information), and (2) acid treatment of mouse splenocytes (RIKEN protocol; Methods and Extended Data Fig. 2b). Seventy-two hours after stress treatment of lung cells, floating spheres appeared amidst cellular debris. Fluorescence microscopy revealed that both *Oct4*-GFP and wild-type spheres emitted low-level broad spectrum fluorescence detectable within both green and red filters, indicating autofluorescence (Fig. 1a). Untreated *Oct4*-GFP ES cells did not emit the same low-level broad spectrum fluorescence as STAP-treated cells. STAP-treated splenocytes formed spheres with lower efficiency, but also appeared autofluorescent.

Flow cytometry indicated STAP-treated *Oct4*-GFP cells did not exhibit *Oct4*-GFP reactivation at levels comparable to control *Oct4*-GFP mouse ES cells, and were indistinguishable from stressed wild-type controls (Fig. 1b). Absence of ES-cell-like levels of *Oct4*, Sox2 and Nanog transcripts and nonspecific immunofluorescence corroborated flow cytometry data (Extended Data Fig. 2c, d). Rare pluripotent cells should generate teratomas in immunocompromised mice^{8,9}, but STAP cells could not, unlike control ES cells (Extended Data Fig. 2e, f). Replication of the poly-L-glycolic acid (PLGA)-based teratoma production method described in the original STAP reports with GFP cells to distinguish host and donor contribution produced distinct masses of connective tissue, muscle and scar, with minimal GFP content, indicating primarily host origin (Fig. 1c, d and Extended Data Fig. 2g). Rare GFP-positive clusters did not form differentiated tissues characteristic of ES-cell-derived teratomas (Fig. 1d). Autofluorescent spheres failed to enter development after

morula aggregation or blastocyst injection (Fig. 1e and Extended Data Fig. 2h–j). Therefore, pluripotency was undetectable in STAP experiments. Six other laboratories (Deng, Hanna, Hochedlinger, Jaenisch, Pei and Wernig) also attempted to generate STAP cells (Table 1) and made the following observations. First, autofluorescent sphere-like aggregates after STAP treatment were universally seen. Second, transgenic reporters used by Obokata and colleagues (GOF18-*Oct4*-GFP, containing the 18-kilobase genomic *Oct4* fragment (GOF18)) and by the Daley, Pei and Hanna laboratories (GOF18-*Oct4*ΔPE-GFP, lacking the *Oct4* proximal enhancer (PE) element) both exhibit activity in pre-implantation embryos, early post-implantation epiblast cells (embryonic day (E) 5.5), germ cells, and mouse ES/iPS cells; however, differential activity in late post-implantation epiblast (E6.5) and early passage mouse epiblast-derived stem cells has been ascribed to the *Oct4* proximal enhancer^{10–12}. Using the same reporter as Obokata and colleagues^{1,2}, the Deng laboratory observed that the GFP signal in chemical iPS cells was easily distinguishable from the autofluorescence of STAP-treated cells (Extended Data Fig. 2k). The Jaenisch, Wernig and Hochedlinger laboratories failed to observe GFP reactivation with *Oct4* or *Nanog* knock-in reporters, excluding a scenario of uncoupling between GFP and endogenous pluripotency expression¹⁰. Despite a range of tested reporters, no group documented authentic *Oct4*/Nanog reporter activation that resembled bona fide ES cells. Third, the Deng laboratory failed to observe *Oct4*, Sox2 and Nanog induction 3 and 7 days after STAP treatment, reducing the likelihood that pluripotency was transiently activated and silenced by day 7 (Extended Data Fig. 2l). Finally, the Hanna, Wernig and Hochedlinger laboratories failed to generate stem-cell lines by culturing STAP-treated cells in leukaemia inhibitory factor (LIF) and adrenocorticotrophic hormone (ACTH)-supplemented medium. In summary, 133 replicate attempts failed to document generation of ES-cell-like cells, corroborating and extending a recent report¹³.

We re-examined the high-throughput sequencing data from the STAP reports to investigate the genetic provenance of parental CD45⁺ cells and converted STAP cells, STAP stem cells and Fgf4-induced stem cells (FI-SCs) (Fig. 1f). Comparative genomic hybridization array data mentioned in the original paper¹ were not publicly released. Copy number variation (CNV) analysis conducted using ChIP-seq input samples revealed a discrepancy in sex across samples as well as chromosomal aberrations (Fig. 1g). In the original STAP reports, the authors stated that they mixed CD45⁺ cells from male and female mice owing to the small number of CD45⁺ cells retrieved from individual neonatal spleens. However, our analysis indicates that CD45⁺ cells were female, whereas the derived cells (STAP cells, STAP stem cells and FI-SCs) were all male, a clear inconsistency. We note that control ES cells were also male (Fig. 1g). FI-SCs possessed trisomy 8, which renders mouse ES cells germline-incompetent¹⁴ (Fig. 1g).

Inferred single nucleotide variants (SNVs) from RNA-seq data allowed classification of samples as genetically similar or dissimilar (Fig. 1h). Control ES cells, parental donor female CD45⁺ cells, STAP cells, and STAP stem cells all possessed similar SNV profiles, consistent with their derivation from a first generation hybrid of C57BL/6/129 strains, the reported genotype (Fig. 1h and Extended Data Fig. 3). By contrast, FI-SCs had an SNV profile that matched a single nucleotide polymorphism (SNP) profile of C57BL/6 strain origin, indicating

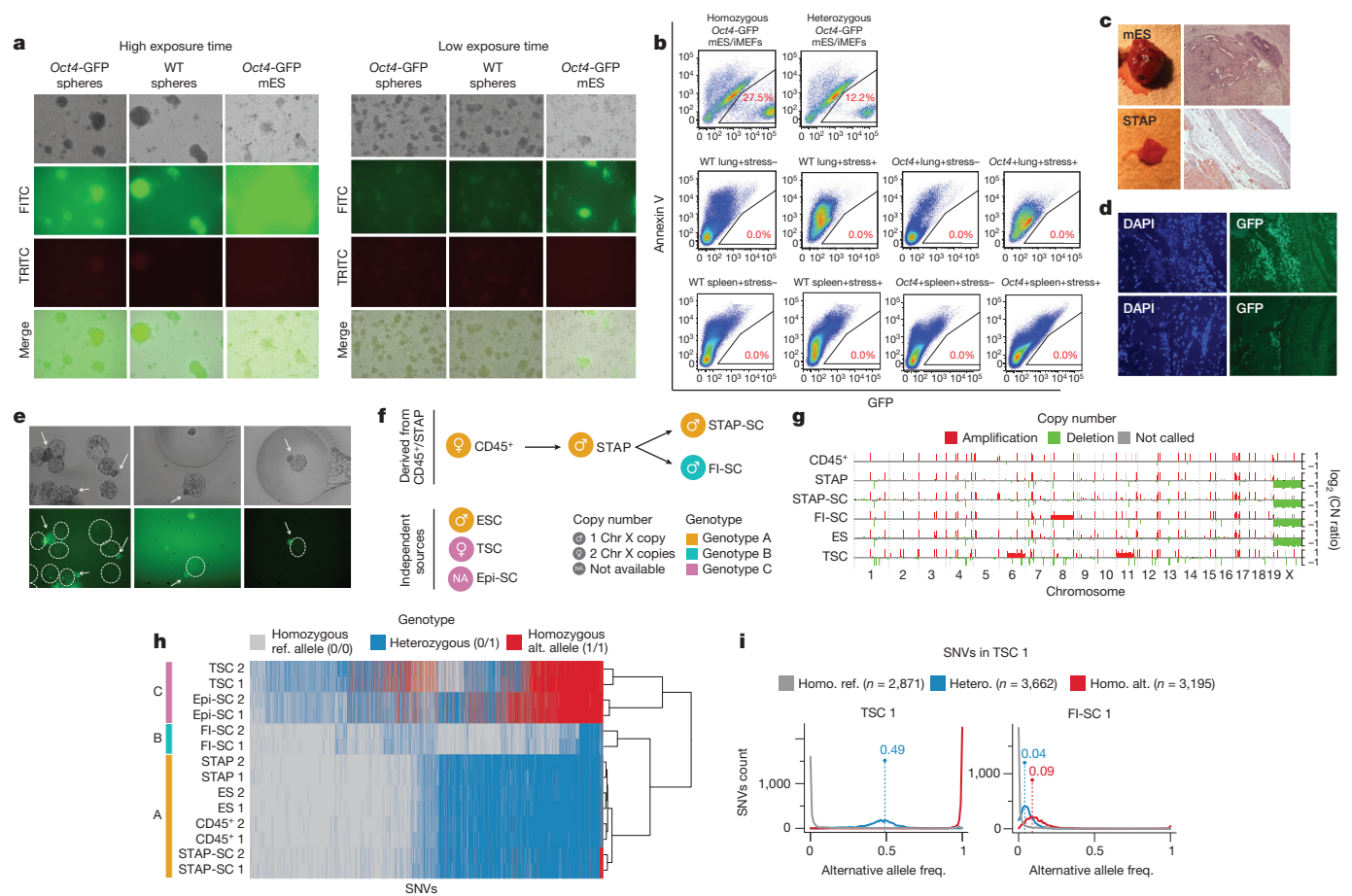


Figure 1 | Characterization of the STAP cell phenomenon. See Supplementary Information for further details. **a**, STAP treatment produces fluorescent signal detected in both FITC (green) and TRITC (red) channels in STAP-treated *Oct4*-GFP and wild-type (WT) cells, consistent with autofluorescence. TRITC signal was not detected in control *Oct4*-GFP mouse ES cells. Note saturation of the green signal in *Oct4*-GFP ES cells at the higher exposure time required to detect FITC from autofluorescent spheres. **b**, Absence of ES-cell-like *Oct4*-GFP reactivation. Representative flow cytometry results 7 days after STAP treatment of lung cells or splenocytes (BWH or RIKEN protocol, respectively) without singlet/doublet exclusion and live/dead-cell discrimination. GFP gates were calibrated based on control *Oct4*-GFP ES cells grown on feeders. Whereas control ES cells are bright and situated at approximately 1×10^5 (arbitrary units), no event resembling *Oct4*-GFP ES cells was detected after STAP treatment. One replicate per protocol is shown. iMEFs, irradiated MEFs. **c**, STAP-treated cells do not form teratomas using PLGA-based teratoma production methods¹. Photograph of control mouse ES-cell-derived teratoma (top left) and non-teratoma STAP-PLGA mass (bottom left). Representative haematoxylin and eosin (H&E) stainings of a control mouse ES-cell-derived teratoma (top right) and the non-teratoma STAP-PLGA mass (bottom right). **d**, STAP-PLGA mixtures present no indication of ES-cell-like *in vivo* differentiation capacity after injection into immunocompromised mice. Note lack of organization into representative tissue structures typically observed in ES-cell-derived teratomas. DAPI, 4',6-diamidino-2-phenylindole. **e**, STAP-treated lung cells fail to incorporate into preimplantation embryos after morula aggregation. **f**, Analysis of sequencing data. Samples are classified based on copy number and genotype. STAP cells, STAP stem cells (STAP-SCs)

and ES cells share similar characteristics for genotype and copy number of chromosome X. **g**, Copy number (CN) profiles, reported as a \log_2 ratio (observed to expected read counts), derived using ChIP-seq input data. Red/green correspond to significant amplifications and deletions ($\log_2(\text{CN}) \geq 0.2$ or ≤ -0.2 and $P \leq 0.01$), respectively. Grey denotes non-significant variants. Note the amplifications of chromosomes 8 (FI-SCs) and 6/11 (TSCs) and the single copy of chromosome X in STAP cells, STAP-SCs, FI-SCs and ES cells. **h**, SNVs inferred from RNA-seq data using the mouse reference genome (derived from C57BL/6 strain). The selected SNVs are classified as homozygous for reference allele (0/0 genotype), homozygous for alternative allele (1/1 genotype) or heterozygous (0/1 genotype). Samples are clustered based on the sum of edit distance between each SNV. Note that each pair of replicates is always grouped together. A subset of samples (CD45⁺, STAP, STAP-SCs and ES cells) shows prevalence of heterozygous alleles (A); FI-SC samples have prevalence of homozygous alleles for the reference variant (B); and, TSC and epiblast stem cell (EpiSC) samples have a larger number of homozygous alternative alleles (C). **i**, Contamination in the FI-SC samples with TSCs. The expected frequency of reads covering the alternative allele for heterozygous SNVs is ~50%, which is observed in all samples including TSCs (left). In FI-SCs, it was ~12% (Extended Data Fig. 3), suggesting false-positive calls or contamination. The alternative allele frequency distributions of TSC homozygous and heterozygous SNVs sets in FI-SCs (right) show peaks at 9% and 4%, respectively. These results indicate that FI-SC samples are approximately 10% contaminated by TSC samples. Original magnifications, $\times 20$ (**a**, **d**, **e**) and $\times 4$ (**c**).

distinct genetic provenance from parental CD45⁺ and STAP samples (Fig. 1h and Extended Data Fig. 3). Independently sourced control epiblast stem cells and trophoblast stem cells (TSCs) had SNV profiles divergent from the CD45⁺ and STAP sample cohort, as expected (Fig. 1h). An anomalous allele frequency distribution observed in

FI-SCs, and reciprocal analyses of FI-SC heterozygous SNVs in TSCs and TSC homozygous and heterozygous SNVs in FI-SCs, revealed that FI-SCs were derived from a C57BL/6 strain origin, with approximately 10% contamination from TSCs (Fig. 1i and Extended Data Fig. 3). These are concordant with the findings from a recent

BRIEF COMMUNICATIONS ARISING

Table 1 | Global efforts to replicate STAP cell generation

Laboratory	Mouse strain	Starting cell types	Pluripotency reporter	STAP generation protocol	Pluripotency assessment assay	Experimental replicates	Conclusion
Jaenisch (MIT, USA)	C57BL6/129; JAX 008214	MEFs; neonatal spleen; neonatal adipocytes; neonatal fibroblasts	<i>Oct4</i> -eGFP knock-in (Jaenisch); <i>Nanog</i> -eGFP knock-in (Jaenisch)	Obokata article ^{1,3} ; Obokata, Sasai & Niwa protocol*	Epifluorescence; FACS	17	Negative
Wernig (Stanford, USA)	C57BL6/129; JAX 008214	MEFs	<i>Oct4</i> -eGFP knock-in (Jaenisch)	Obokata article ^{1,3}	Epifluorescence; FACS; STAP-SC	3	Negative
Pei (Guangzhou, China)	C57BL6/CBA; JAX 004654	Spleen MNCs, CD45 ⁺ spleen cells, MEFs, pre-iPS cells, EpiSCs, mammary epithelial cells	GOF18-Oct4ΔPE-GFP transgene (Scholer/Mann)	Obokata article ^{1,3} ; Obokata, Sasai & Niwa protocol*	Epifluorescence; FACS	11	Negative
Deng (Peking, China)	C57BL6 × ICR	MEFs, neonatal fibroblasts, neonatal bone marrow, neonatal brain, neonatal cardiac muscle cells, neonatal heart, neonatal lung, neonatal spleen, adult spleen	GOF18-Oct4-GFP transgene (Scholer; identical to Obokata article reporter ^{1,3})	Obokata article ^{1,3} ; Obokata, Sasai & Niwa protocol*	Epifluorescence; FACS; qPCR	37	Negative
Hanna (Weizmann, Israel)	C57BL6/CBA; JAX 004654	Neonatal CD45 ⁺ spleen cells, neonatal liver, neonatal brain	GOF18-Oct4ΔPE-GFP transgene (Scholer/Mann)	Obokata article ^{1,3}	Epifluorescence; FACS; STAP-SC	5	Negative
Hochedlinger (Harvard, USA)	C57BL6/129; JAX 008214	Neonatal CD45 ⁺ splenocytes, E14.5 MEFs	<i>Oct4</i> -eGFP knock-in (Jaenisch)	Obokata article ^{1,3} ; BWH protocol†; Obokata, Sasai & Niwa protocol*	Epifluorescence; immunofluorescence; FACS; STAP-SC	6	Negative
Daley (Harvard, USA)	C57BL6/CBA; JAX 004654	MEFs, neonatal CD45 ⁺ spleen cells, neonatal lung cells, neonatal liver, neonatal heart, neonatal brain	GOF18-Oct4ΔPE-GFP transgene (Scholer/Mann)	Obokata article ^{1,3} ; BWH protocol†; Obokata, Sasai & Niwa protocol*	Epifluorescence; FACS; qPCR; immunofluorescence; teratoma; chimaera	54	Negative

The STAP papers^{1,2} were published in *Nature* on 29 January 2014 (both retracted^{3,4} in July 2014). Below, we provide the time period in which laboratories attempted to replicate STAP cell generation. The Jaenisch laboratory performed STAP replication experiments from February to April 2014. The Hanna laboratory performed STAP replication experiments from February to March 2014. The Wernig laboratory performed STAP replication experiments in February 2014. The Deng laboratory performed STAP replication experiments from February to July 2014. The Pei laboratory performed STAP replication experiments from February to May 2014. The Hochedlinger laboratory started STAP replication experiments on the day of STAP paper publication (29 January 2014) and continued to April 2014. The Daley laboratory performed STAP replication experiments and characterization of the STAP phenomenon from February to November 2014.

*Obokata, Sasai & Niwa protocol: <http://dx.doi.org/10.1038/protex.2014.008>

†BWH protocol: see Supplementary Information.

eGFP, enhanced GFP; EpiSCs, epiblast stem cells; MEFs, mouse embryonic fibroblasts; MNCs, mononuclear cells; qPCR, quantitative PCR; STAP-SC, STAP stem cells.

RIKEN report (<http://www3.riken.jp/stap/e/c13document52.pdf>). This contamination with TSCs explains the high-grade placenta-forming capacity reported for the FI-SCs², an unusual feature that implied totipotency, but which seems to have been due to admixture of cells.

In summary, our replication attempts and genetic analysis indicate that existing STAP protocols are neither robust nor reproducible. To substantiate future claims of reprogramming and alternative states of potency, we urge a rigorous application of several independent means for validating functional pluripotency and genomic profiling to confirm cell line provenance. Ultimately, the essential standard of robustness and reproducibility must be met for new claims to exert a positive and lasting influence on the research community.

Alejandro De Los Angeles^{1,2,3*}, Francesco Ferrari^{4,†*}, Yuko Fujiwara^{1,2,3}, Ronald Mathieu⁵, Soohyun Lee⁴, Semin Lee⁴, Ho-Chou Tu^{1,2,3}, Samantha Ross^{1,2,3}, Stephanie Chou^{1,2,3}, Minh Nguyen^{1,2,3}, Zhaoting Wu^{1,2,3}, Thorold W. Theunissen⁶, Benjamin E. Powell⁶, Sumeth Imsoonthornruksa^{6,†}, Jiekai Chen⁷, Marti Borkent^{2,3,8,9}, Vladislav Krupalnik¹⁰, Ernesto Lujan¹¹, Marius Wernig¹¹, Jacob H. Hanna¹⁰, Konrad Hochedlinger^{2,3,8}, Duanqing Pei⁷, Rudolf Jaenisch⁶, Hongkui Deng¹², Stuart H. Orkin^{1,2,3}, Peter J. Park⁴ & George Q. Daley^{1,2,3}

¹Stem Cell Transplantation Program, Division of Pediatric Hematology Oncology, Children's Hospital Boston, and Dana-Farber Cancer Institute, Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, Massachusetts 02115, USA.

²Harvard Stem Cell Institute, Cambridge, Massachusetts 02138, USA.

³Howard Hughes Medical Institute, Boston, Massachusetts 02115, USA.

⁴Department of Biomedical Informatics, Harvard Medical School, Boston, Massachusetts 02115, USA.

email: peter_park@hms.harvard.edu

⁵Children's Hospital & Harvard Stem Cell Institute Flow Cytometry Research Facility, Children's Hospital Boston, Boston, Massachusetts 02115, USA.

⁶Whitehead Institute for Biomedical Research, Cambridge, Massachusetts 02142, USA.

⁷South China Institute for Stem Cell Biology and Regenerative Medicine, Guangzhou Institutes of Biomedicine and Health, Chinese Academy of Sciences, Guangzhou, 510530, China.

⁸Massachusetts General Hospital Cancer Center and Center for Regenerative Medicine, Boston, Massachusetts 02114, USA.

⁹Department of Development and Reproduction, Erasmus Medical Center, 3015 CN Rotterdam, Netherlands.

¹⁰Department of Molecular Genetics, Weizmann Institute of Science, Rehovot 76100, Israel.

¹¹Institute for Stem Cell Biology and Regenerative Medicine, Stanford University, Stanford, California 94305, USA.

¹²College of Life Sciences and Peking-Tsinghua Center for Life Sciences, Peking University, Beijing 100871, China.

†Present addresses: IFOM, The FIRC Institute of Molecular Oncology, Milan 20139, Italy (F.F.); Center for Biomolecular Structure, Function and Application Research Unit and School of Biotechnology, Institute of Agricultural Technology, Suranaree University of Technology, Nakhon Ratchasima 30000, Thailand (S.I.).

*These authors contributed equally to this work.

Received 10 November 2014; accepted 22 July 2015.

- Obokata, H. *et al.* Stimulus-triggered fate conversion of somatic cells into pluripotency. *Nature* **505**, 641–647 (2014).

2. Obokata, H. *et al.* Bidirectional potential in reprogrammed cells with acquired pluripotency. *Nature* **505**, 676–680 (2014).
3. Obokata, H. *et al.* Retraction: stimulus-triggered fate conversion of somatic cells into pluripotency. *Nature* **511**, 112 (2014).
4. Obokata, H. *et al.* Retraction: bidirectional potential in reprogrammed cells with acquired pluripotency. *Nature* **511**, 112 (2014).
5. Okita, K., Ichisaka, T. & Yamanaka, S. Generation of germline-competent induced pluripotent stem cells. *Nature* **448**, 313–317 (2007).
6. Nichols, J. *et al.* Formation of pluripotent stem cells in the mammalian embryo depends on the POU transcription factor Oct4. *Cell* **95**, 379–391 (1998).
7. Szabó, P. E., Hubner, K., Scholer, H. & Mann, J. R. Allele-specific expression of imprinted genes in mouse migratory primordial germ cells. *Mech. Dev.* **115**, 157–160 (2002).
8. Lawrenz, B. *et al.* Highly sensitive biosafety model for stem cell-derived grafts. *Cytotherapy* **6**, 212–222 (2004).
9. Cao, F. *et al.* Spatial and temporal kinetics of teratoma formation from mouse embryonic stem cell transplantation. *Stem Cells Dev.* **16**, 883–892 (2007).
10. Han, D. W. *et al.* Epiblast stem cell subpopulations represent mouse embryos of distinct pregastrulation stages. *Cell* **143**, 617–627 (2010).
11. Yeom, Y. I. *et al.* Germline regulatory element of Oct4 specific for the totipotent cycle of embryonal cells. *Development* **122**, 881–894 (1996).
12. Ohbo, K. *et al.* Identification and characterization of stem cells in prepubertal spermatogenesis in mice. *Dev. Biol.* **258**, 209–225 (2003).
13. Tang, M. K. *et al.* Transient acid treatment cannot induce neonatal somatic cells to become pluripotent stem cells. *F1000Res.* **3**, 102 (2014).
14. Kim, Y. M., Lee, J. Y., Xia, L., Mulvihill, J. J. & Li, S. Trisomy 8: a common finding in mouse embryonic stem (ES) cell lines. *Mol. Cytogenet.* **6**, 3 (2013).
15. Hou, P. *et al.* Pluripotent stem cells induced from mouse somatic cells by small-molecule compounds. *Science* **341**, 651–654 (2013).

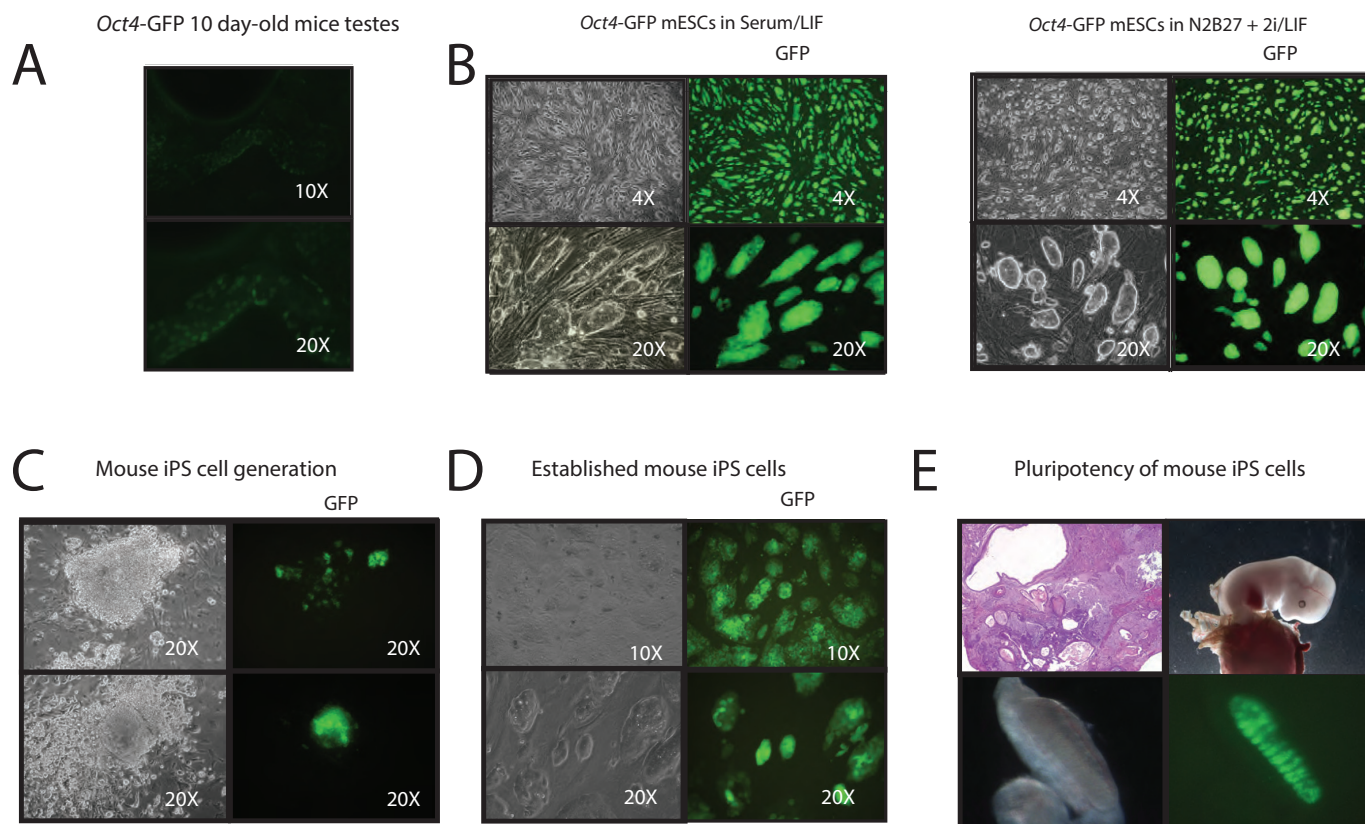
Supplementary Information is available in the online version of the paper.

Author Contributions A.D.L.A. and F.F. contributed equally to this work. A.D.L.A. and G.Q.D. conceived and designed the project. A.D.L.A. performed experiments within the Vacanti laboratory, where indicated. A.D.L.A., Y.F., R.M., H.-C.T., S.C. and Z.W. analysed STAP experiments. S.R. facilitated teratoma injections. T.W.T., B.E.P., S.I., J.C., M.B., V.K., E.L., M.W., J.H.H., K.H., D.P., R.J. and H.D. contributed STAP replication data. F.F. performed bioinformatics analyses, assisted by So.L. and Se.L., and supervised by P.J.P. A.D.L.A., F.F., P.J.P. and G.Q.D. wrote the manuscript.

Competing Financial Interests Declared none.

doi:10.1038/nature15513

BRIEF COMMUNICATIONS ARISING

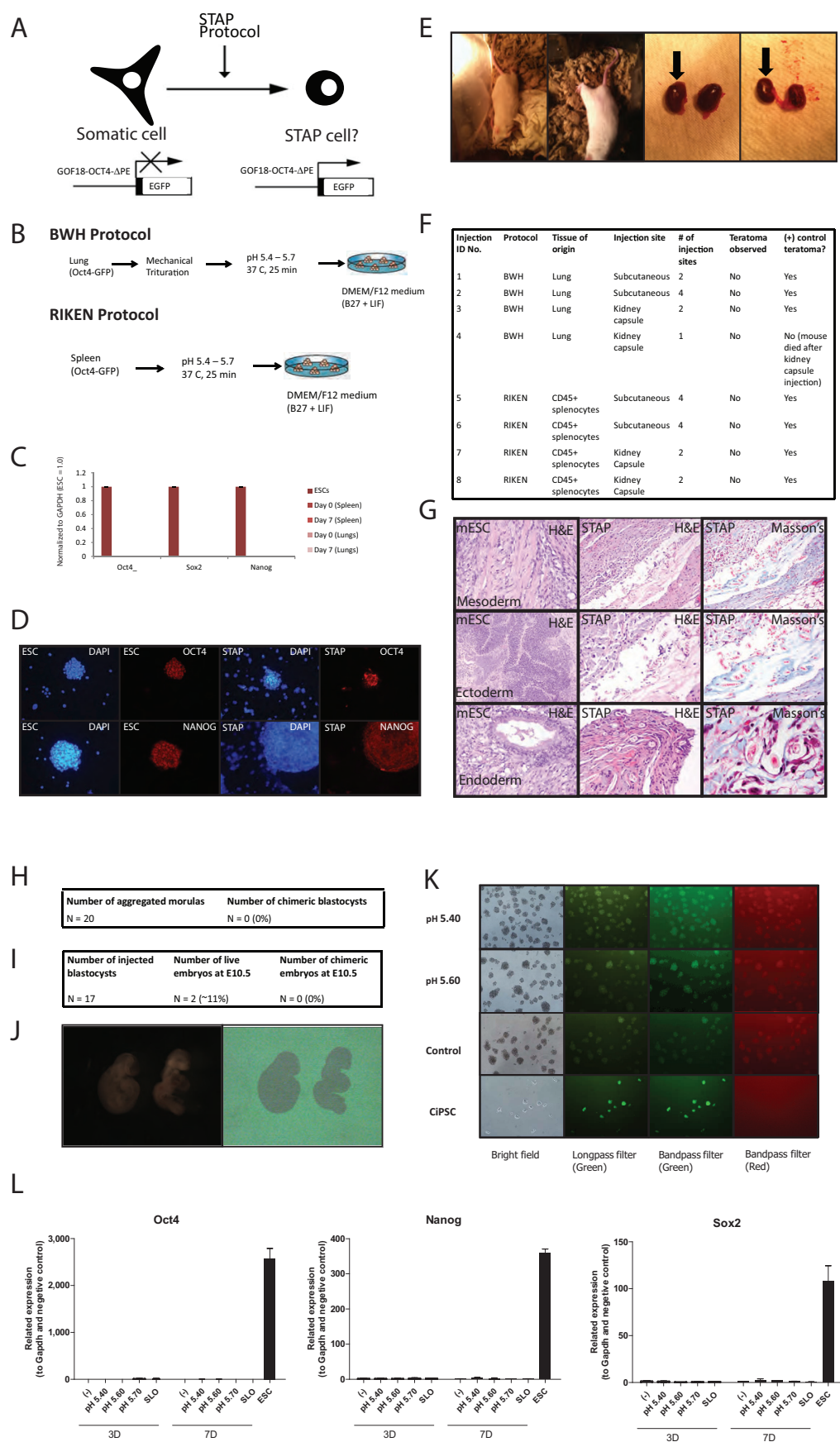


Extended Data Figure 1 | Validation of the *Oct4*-GFP transgenic reporter.

a, Context-appropriate expression of the GOF18-*Oct4*ΔPE-GFP transgene reporter in the testes of 10-day-old neonatal male mice. **b**, STAP replication culture reagents sustain *Oct4*-GFP signal in *Oct4*-GFP mouse ES cells. Vacanti laboratory LIF and B27 supplement sustain self-renewal and strong GFP signal of *Oct4*-GFP mouse ES cells in serum/LIF (left) and N2B27 minimal media (see Methods) plus 2i/LIF (MEK inhibitor PD0325901 and GSK3-β inhibitor CHIR99021) (right). **c**, Reactivation of the GOF18-*Oct4*ΔPE-GFP reporter during direct reprogramming of MEFs by Oct4, Sox2 and Klf4. Left, phase-contrast images of founder GOF18-mouse iPS cells. Right, GFP signal in primary GOF18-mouse iPS cells. Note the heterogeneous reactivation of the GOF18-*Oct4*ΔPE-GFP reporter in primary founder mouse iPS cell colonies (derived in knockout serum replacement/LIF). **d**, GOF18-*Oct4*ΔPE-GFP reporter expression in established mouse iPS cell lines (passage 12). Left, phase-contrast images of established GOF18-mouse iPS cells. Mouse iPS cells were

maintained on feeders in serum/LIF media. Right, note GFP signal in GOF18-mouse iPS cells. GFP is observed in essentially all iPS cell colonies and in most cells in each colony. GFP heterogeneity was slightly increased in GOF18-iPS cells compared with GOF18-ES cells. **e**, Developmental potential of GOF18-iPS cells. Top left, phase-contrast image of a teratoma generated from GOF18-iPS cells. Original magnification, ×4. Top right, to assess the developmental potential of GOF18-iPS cells, 'all iPS cell embryos' were generated by injection of GOF18-iPS cells into 4N blastocysts ('tetraploid complementation'). A photograph of a live E13.5 embryo generated from GOF18-iPS cells is shown. Bottom row, gonadal contribution in all-iPS-cell embryos indicates GOF18-iPS cells are highly pluripotent. GFP is expressed in E13.5 days post-coitum (dpc) male gonads, and fluorescent cords are visible. The silencing of GFP in surrounding cells re-confirms the context-appropriate expression of the *Oct4*-GFP reporter.

BRIEF COMMUNICATIONS ARISING

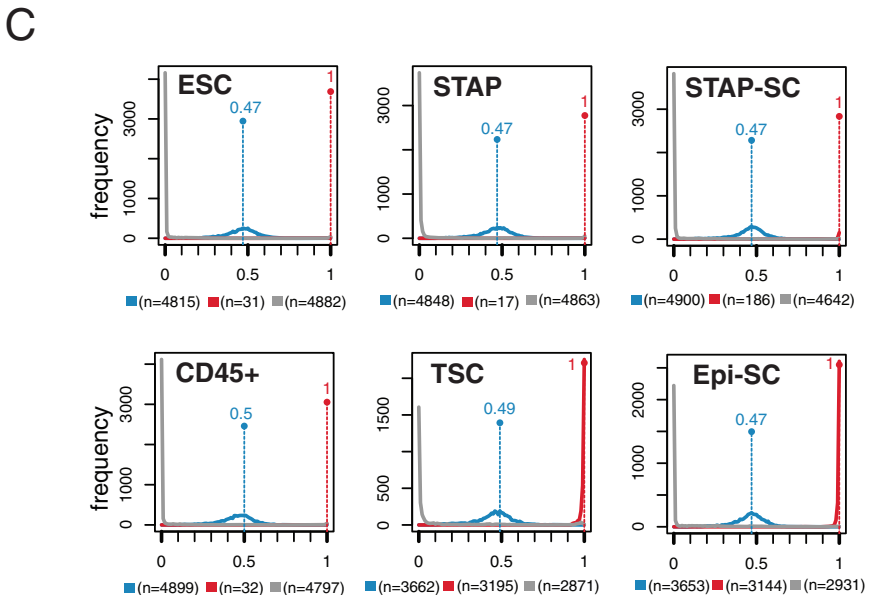
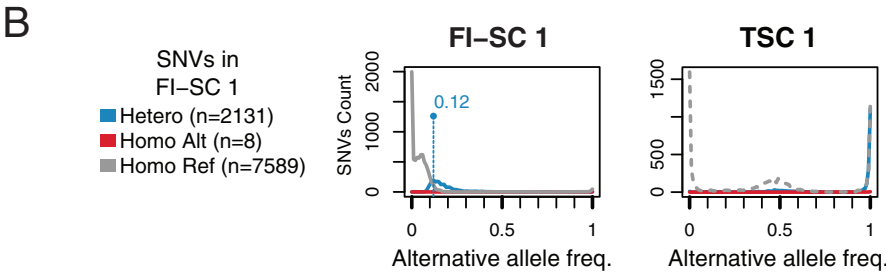
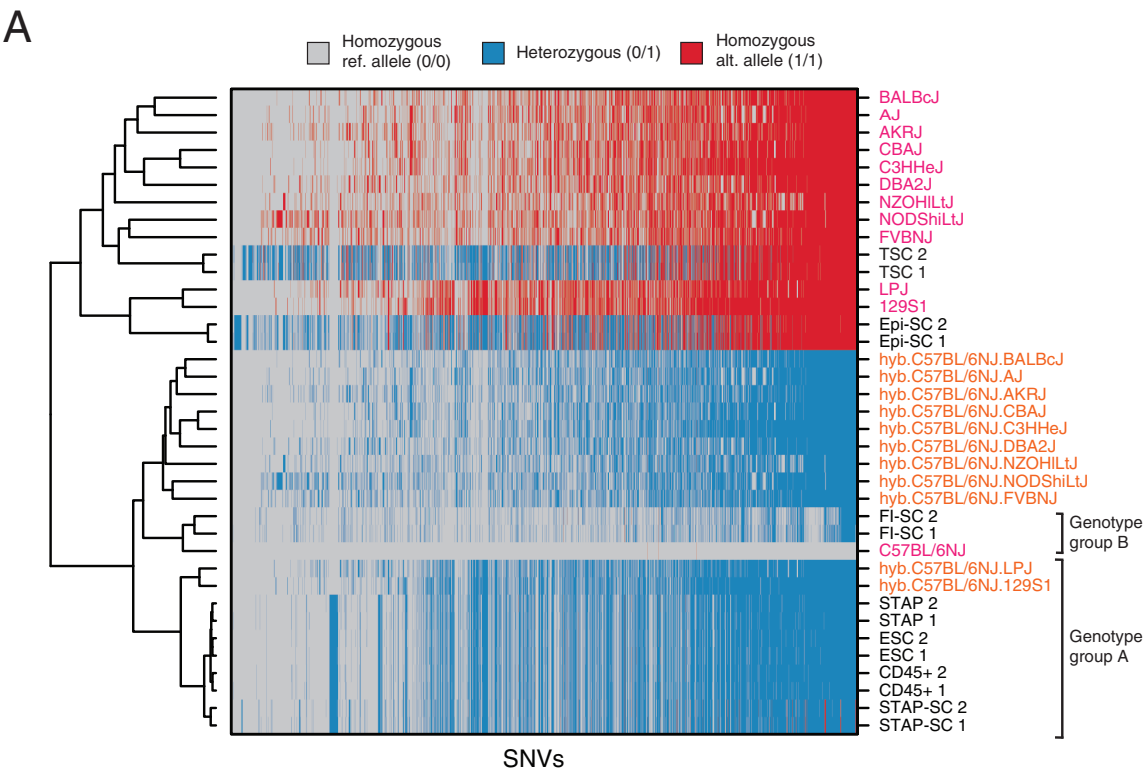


BRIEF COMMUNICATIONS ARISING

Extended Data Figure 2 | STAP replication data. **a**, Experimental scheme. Reactivation of the transgenic GOF18-Oct4ΔPE-GFP (*Oct4*-GFP) reporter to detect reprogramming after STAP treatment of somatic cells. **b**, Two STAP protocol variants. BWH: mechanical trituration and low pH treatment of lung cells. RIKEN: low pH treatment of spleen cells. Stressed cells were plated into non-adhesive dishes and cultured in DMEM/F12 medium plus B27 and LIF. **c**, qPCR analysis 7 days after STAP induction. Expression levels of *Oct4*, *Sox2* and *Nanog* transcripts at ES-cell-like levels were not observed in lung cells or splenocytes after treatment with the BWH or RIKEN STAP protocol, respectively. Levels normalized to *Gapdh*. One replicate per protocol is shown. Although low level *Sox2* and *Nanog* upregulation (1–2 ΔC_t cycles; data not shown) was inconsistently observed, we speculate that minimal induction of *Sox2* and *Nanog* messenger RNA may be due to relaxed transcriptional control in stressed cells. **d**, Nonspecific staining observed in STAP-treated cells suggests immunofluorescence artefacts. ES cells and autofluorescent spheres (BWH protocol) were processed in parallel and stained with Oct4 and Nanog antibodies. In contrast to the specific nuclear signal observed in positive-control ES cells, nonspecific and non-nuclear staining is observed in spheres generated after STAP treatment. Original magnification, $\times 20$. **e**, Assessing the presence of rare ES-cell-like cells in STAP-treated cultures by teratoma formation assays. STAP-treated cells were transplanted subcutaneously or into the kidney capsule to detect rare ES-cell-like pluripotent cells. If ES-cell-like cells are generated after transient low pH treatment with/without mechanical trituration, a teratoma containing elements of all three germ layers should form. STAP-treated cells did not form teratomas using conventional teratoma generation protocols. Left two images, immunocompromised mice injected subcutaneously with STAP-treated cells, which do not exhibit teratoma-like mass formation after approximately 4 months of observation. Right two images, kidneys after STAP-treated cells were transplanted into the kidney capsule indicating lack of teratoma-like formation after 3 months of observation. Black arrows indicate kidney transplanted with STAP-treated cells; second kidney from same mouse not transplanted with STAP-treated cells. **f**, Immunocompromised NOD/SCID mice transplanted with STAP-treated cells did not form teratoma-like masses. Summary of teratoma injection experiments. Every assessable injection of mouse ES cells produced teratomas (7 out of 8 positive-control ES-cell-injected mice formed teratomas within 3–4 weeks. The mouse that did not form a teratoma immediately died from surgical complications and therefore was discarded from the analysis). $n = 8$ independent injection sessions; $n = 21$ injection sites. Therefore, STAP-treated cells did not form teratomas using conventional methods. **g**, Extended histological analysis of a recovered STAP-PLGA mass (as in Fig. 1c). Obokata and colleagues^{1,2} reported a distinct teratoma production method that involved seeding STAP-treated cells onto a PLGA scaffold before implantation into

immunocompromised mice. Around 10–20 million STAP-treated cells from GFP-positive mice were seeded into PLGA. GFP-positive cells were used to distinguish donor- and host-derived tissues. Left, positive-control ES cells formed teratomas with tissue derivatives of all three germ layers. Left, original magnifications (from top to bottom): $\times 40$, $\times 20$, $\times 40$. Middle, recovered STAP-PLGA mass, H&E staining. Middle, original magnifications (from top to bottom): $\times 20$, $\times 40$, $\times 40$. Right, recovered STAP-PLGA mass, Masson's staining (used to illustrate collagen deposition or presence of an inflammatory reaction, which commonly occur in response to foreign body implants). Right, original magnifications (from top to bottom): $\times 20$, $\times 40$, $\times 60$. All images were obtained from formalin-fixed/paraffin-embedded tissue sections. STAP-treated autofluorescent spheres failed to re-enter development after morula aggregation. Unlike ES or iPS cells, autofluorescent spheres failed to incorporate into the inner cell mass of the host embryos ($n = 20$), suggesting incompatibility with the pre-implantation embryo. **i**, STAP-treated autofluorescent spheres failed to re-enter development after blastocyst injection. Mechanically disaggregated autofluorescent spheres were injected into pre-implantation blastocysts and implanted into pseudopregnant mice. From 17 implanted embryos, only two were recovered, which were developmentally abnormal, suggesting that the other 15 embryos died or were resorbed. **j**, Contribution of STAP-treated lung cells to chimaeras was not detected after blastocyst injection. Images of two abnormal E10.5 embryos with no obvious GFP signal that would indicate integration of donor test cells into the developing host embryo. Original magnification, $\times 10$. **k**, Autofluorescence and *Oct4*-GFP fluorescence were distinguishable by fluorescence microscopy in cells containing the same *Oct4*-GFP transgenic reporter used by Obokata and colleagues^{1,2} (data from Deng laboratory). MEFs with the same transgenic *Oct4*-GFP reporter (GOF18-Oct4-GFP, intact PE) used in ref. 1 (passage 0) were treated with low pH solutions (pH 5.4 and 5.6, respectively). MEFs without low pH treatment were used as a negative control. After treatment, samples were cultured in suspension. Chemically induced pluripotent stem cells (CiPSC)¹⁵ containing the transgenic *Oct4*-GFP reporter were used as a positive control for green fluorescence. GFP fluorescence was detected using a long-pass and band-pass filter. Red fluorescence was also observed in low-pH-treated MEFs, but not in CiPSC, as shown in the right column. Scale bar, 100 μm . **l**, ES-cell-like levels of *Oct4*, *Sox2* and *Nanog* mRNA (analysed by qPCR) were not observed 3 days after STAP treatment of MEFs (data from Deng laboratory). MEFs were treated with low pH solutions and cultured in suspension for 3 and 7 days (following the RIKEN STAP protocol) and analysed. R1 ES cells were used as a positive control. MEFs that were not subjected to the RIKEN STAP protocol but cultured in suspension medium were used as the negative control (–).

BRIEF COMMUNICATIONS ARISING



BRIEF COMMUNICATIONS ARISING

Extended Data Figure 3 | Re-analysis of published STAP RNA-seq data.

a. SNVs inferred from RNA-seq (from Fig. 1h) were further filtered to select only the known SNPs across mouse strains based on the Sanger database (see Methods). We compared the SNV profile inferred from STAP RNA-seq data to the expected profiles (simulated based on the known SNPs) in different mouse strains (magenta) as well as simulated first-generation hybrids of C57BL/6NJ and each of the other strains considered (orange). Selected SNVs are classified as homozygous for reference allele (0/0 genotype), homozygous for alternative allele (1/1 genotype) or heterozygous (0/1 genotype) at each locus. Samples are clustered based on the sum of edit distance between each SNV using complete linkage hierarchical clustering. Note that replicates of the same experiment are always grouped together. A subset of samples (CD45⁺, STAP, STAP stem cells and ES cells) (genotype A as in Fig. 1h) are clustered with simulated first-generation hybrids of C57BL/6NJ and 129S1, in accordance with Obokata *et al.*^{1,2} (LPJ strains have a profile similar to 129S1 for the selected SNVs). Whereas FI-SCs (genotype B as in Fig. 1h) are closer to the C57BL/6NJ strain (not the hybrid), EpiSC samples cluster with 129S1 or LPJ simulated SNVs profiles, both with some differences. Again the high similarity between 129S1 and LPJ for these selected SNVs does not allow

discriminating which of them is closer to EpiSC samples. Finally, TSC samples are clustered with other strains not mentioned by Obokata *et al.*^{1,2}. Overall, it is clear that TSC (as well as EpiSC) samples are derived from independent sources compared with STAP cells. **b.** Allele frequency distribution for SNVs shows number of reads for alternative alleles compared to the total number of reads for each SNV. The frequency of reads covering the alternative allele for heterozygous SNVs is expected to be approximately 50%, but in FI-SCs, it is nearly 12% (left, blue), suggesting false-positive calls or contamination (default thresholds in the variant calling algorithm result in incorrect classification of calls). We found that these FI-SC 'heterozygous' SNVs are predominantly homozygous for the alternative allele in TSCs (right, blue line), suggesting TSC samples as a contamination source in FI-SCs. The additional plots in Fig. 1i confirm that FI-SC samples are approximately 10% contaminated by TSC samples. **c.** Allele frequency distributions were independently calculated for all samples. As expected, the frequency of reads covering the alternative allele for heterozygous SNVs is approximately 50% (blue line) in all samples except FI-SCs (see **b**). In these plots, the first replicate (replicate 1) for each RNA-seq sample is reported; an almost identical profile is observed in each replicate pair.

FORUM Parkinson's disease

Crystals of a toxic core

An ultra-high-resolution structure of the core segment of assembled α -synuclein — the protein that aggregates in the brains of patients with Parkinson's disease — has been determined. A neurobiologist and a structural biologist discuss the implications of this advance. [SEE ARTICLE P.486](#)

THE PAPER IN BRIEF

- A small segment of α -synuclein is thought to form the core of the protein fibrils that are associated with Parkinson's disease and other synucleinopathies.
- On page 486 of this issue, Rodriguez *et al.*¹ used the sophisticated electron-diffraction technique MicroED to determine the structure of this

tiny core segment of just 11 amino-acid residues.

- The 1.4-ångström structure is the highest resolution yet achieved through cryo-electron-microscopy methods.
- The authors also present a MicroED structure of a segment of assembled α -synuclein that is mutated in some cases of familial Parkinson's disease.

Fibril features

MICHEL GOEDERT

The abnormal assembly of α -synuclein is central to Parkinson's disease². Fibrils formed from α -synuclein (Lewy pathology) are seen in some brain neurons of more than 95% of patients with the disease, and their formation is associated with neurodegeneration³. Certain mutations in the α -synuclein gene, SNCA, and multiplications thereof, cause rare cases of Parkinson's disease. Sequence variants in the gene's regulatory region are associated with increased disease risk. Moreover, overexpression of mutant human α -synuclein in animal models causes its aggregation and neurodegeneration.

α -Synuclein is a 140-amino-acid protein that is abundant in nerve cells, where it is concentrated in nerve terminals. The protein binds to lipids through its amino-terminal half, which comprises seven imperfect repeat sequences. Upon lipid binding, α -synuclein takes on a partly α -helical structure. Under pathological conditions, it self-assembles into oligomers and fibrils. A seed of assembled α -synuclein can trigger aggregation of the soluble protein, and these insoluble aggregates slowly propagate through the brain. The long interval between the formation of the first protein inclusions and the appearance of disease symptoms opens a therapeutic window, provided that sufficiently sensitive diagnostic techniques can be developed.

Unbranched α -synuclein fibrils are 5–10 nanometres in diameter and up to several

micrometres long. They assemble from the full-length protein, but only approximately amino acids 30–100 make up the structured part. Mutations known to cause Parkinson's disease are located between residues 30 and 53. Like other insoluble fibrous protein aggregates (known as amyloids), α -synuclein fibrils contain linked β -sheet structures.

Rodriguez and colleagues present structures of an 11-amino-acid peptide corresponding to residues 68–78 of α -synuclein (Fig. 1), and of a peptide of residues 47–56 that contains the disease-causing mutation A53T. What evidence

The authors postulate that this mutated region interacts with the fibril-forming core to enhance aggregation.

is there that residues 68–78 form the core of α -synuclein fibrils? Previous work has shown that deletion of residues 71–82 abolishes the ability of α -synuclein to assemble into fibrils, propagate and be neurotoxic, and that a peptide of these amino acids will assemble into fibrils. Similarly, deletion of residues 66–74 abolishes assembly and this peptide can also form fibrils. Residues 68–78, studied by Rodriguez *et al.*, can assemble into fibrils too⁴, and the electron-diffraction pattern produced by these assemblies resembles that of fibrils made from full-length α -synuclein. The authors show that fibrils of peptide 68–78 are toxic when externally applied to cells from tumour lines. However, aggregates of α -synuclein form inside cells in Parkinson's disease, so the relevance of this extracellular toxicity is not clear.

Previous studies of fibrils assembled from

full-length α -synuclein have shown that residues 68–78 make up one of several β -strands. Although a complete description of the fibril will require the atomic structures of all β -strands and the regions in between, Rodriguez and colleagues' findings echo previous atomic structures of amyloid-forming peptides⁵. The structures show paired β -sheets with parallel β -strands in each sheet and antiparallel β -strands between sheets (Fig. 1). However, the zipper structure that marks the region between the paired sheets is longer than in other structures, and each pair of β -sheets contains two water molecules, instead of being dry. This new structural information may contribute to the development of molecules that can inhibit the formation of α -synuclein fibrils, as has been shown for the aggregates of tau proteins associated with Alzheimer's disease⁶.

The structure of the assembled mutated A53T α -synuclein peptide shows pairs of interdigitating β -sheets, but with β -strands kinked at residue 51. The authors postulate that this mutated region interacts with the fibril-forming core to enhance aggregation. It remains to be seen whether this model can account for the effects of other mutations in this region that cause Parkinson's disease. Furthermore, the fact that the mutations A30P and E46K, which also cause Parkinson's disease, lie outside the regions studied, suggests that further structural surprises may be in store.

Michel Goedert is at the MRC Laboratory of Molecular Biology, Cambridge CB2 0QH, UK. e-mail: mg@mrc-lmb.cam.ac.uk

Electron diffraction

YIFAN CHENG

The toxic cores of α -synuclein form well-ordered three-dimensional crystals (Fig. 1). But these crystals are so small that they are invisible by light microscopy, and are thus not amenable to structure determination

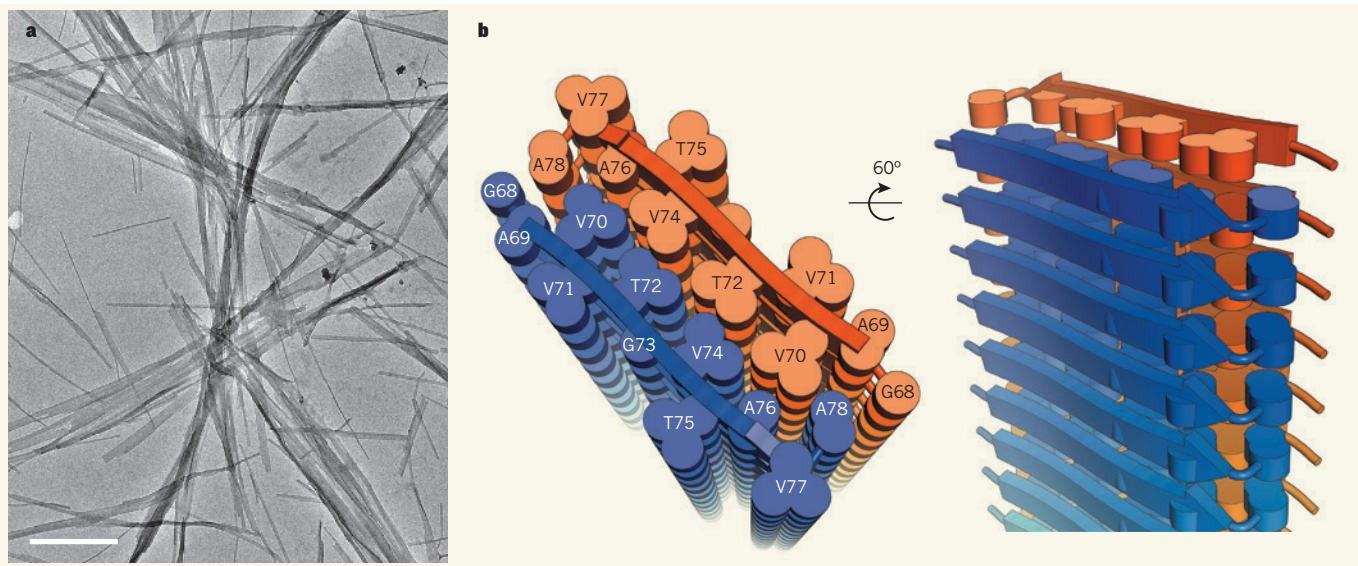


Figure 1 | The core of α -synuclein fibrils. An 11-amino-acid segment of α -synuclein is thought to form the core of the protein fibrils that are seen in patients with Parkinson's disease. **a**, Crystals of this core are so small that they can be seen only by electron microscopy. (Scale bar, 600 nm.) **b**, Rodriguez and colleagues' MicroED structure of the crystal¹ reveals that each peptide forms a β -strand, and these strands pair and stack together to form β -sheets. The structure also reveals two water molecules (not shown) between the paired β -sheets. (Amino-acid residues 68–78 in each strand are labelled.)

using conventional X-ray crystallographic techniques, or even X-ray free-electron laser (XFEL) technology. Now, Rodriguez *et al.* have solved the crystals' structure using MicroED — a method based on cryo-electron microscopy (cryo-EM).

Cryo-EM encompasses several techniques used in structural biology. Single-particle cryo-EM, which determines structures by averaging images of many individual molecules, has already produced high-resolution structures of molecules that were for decades beyond the reach of other crystallographic methods⁷. Electron crystallography has also produced atomic structures from molecules that form 2D crystals. The development of MicroED, which uses electron diffraction to determine the structure of microscopic 3D crystals⁸, has added a new method to the cryo-EM repertoire.

Electron diffraction is an established EM technique that has been used to determine the atomic structures of membrane proteins, such as water channels, that form well-ordered 2D crystals⁹. But for 3D crystals, the situation is more complicated. The primary concerns are dynamic (multiple) scattering, and an inability to index and merge diffraction patterns collected from crystals that vary in size and morphology. MicroED resolves the indexing problem by tilting the crystal in the electron microscope and collecting multiple diffraction patterns from a single crystal, in much the same way as in X-ray crystallography. Continuous rotation of the crystals during data collection also attenuates the dynamic-scattering effect¹⁰.

Rodriguez and colleagues' structures are the first to be determined by MicroED from a molecule of previously unknown structure. They are also the highest-resolution structures

determined using any cryo-EM technique. The study thus demonstrates the tremendous potential of MicroED for use in cases where other crystallographic methods cannot be used. The interaction between an electron beam and a specimen is much stronger than with X-rays, enabling the collection of high-quality diffraction data from tiny crystals. As predicted more than 15 years ago¹¹, the charge of an electron makes it relatively easy for electron diffraction to visualize charged atoms in a structure, such as protons, which require high resolution to be resolved by X-ray crystallography. The instrumentation is readily available,

The structures are the first to be determined by MicroED from a molecule of previously unknown structure.

in that an electron microscope can be operated either as a microscope to produce an image of the specimen, such as for single-particle cryo-EM, or as a diffractometer to produce a diffraction pattern of the specimen. The data-collection and processing methods involved in MicroED are similar to those used in X-ray crystallography and, in comparison with XFELs, the instrumentation cost is amazingly low and accessibility substantially greater.

MicroED does have its limitations. The 'phase problem' of crystallography — the fact that the phases of diffractions cannot be measured — can be particularly challenging in this approach. It is not easy to change the wavelength of an electron beam, so the technique of using multi-wavelength anomalous diffraction to determine phase will probably not be applicable. And it is not clear whether isomorphous replacement, in which heavy

metal atoms are inserted into the structure, will work for *ab initio* phasing, because dynamic scattering may reduce the diffraction signals generated by heavy metals. So far, all structures determined by MicroED used the molecular-replacement method for phase determination, but it remains to be seen how the phasing problem will be resolved in future studies.

There is probably also a constraint on the size of crystals that can be studied by MicroED. The strong scattering makes large crystals impenetrable by electron beams, and merging diffraction patterns from different crystals may be difficult with crystals that have only a few unit cells in one direction. Nonetheless, MicroED provides another highly promising and complementary tool for structural biologists. ■

Yifan Cheng is at the Howard Hughes Medical Institute and the Department of Biochemistry and Biophysics, University of California, San Francisco, San Francisco, California 94143, USA.
e-mail: ycheng@ucsf.edu

- Rodriguez, J. A. *et al.* *Nature* **525**, 486–490 (2015).
- Goedert, M., Spillantini, M. G., Del Tredici, K. & Braak, H. *Nature Rev. Neurol.* **9**, 13–24 (2013).
- Osterberg, V. R. *et al.* *Cell Rep.* **10**, 1252–1260 (2015).
- El-Agnaf, O. M. A. & Irvine, B. G. *Biochem. Soc. Trans.* **30**, 559–565 (2002).
- Sawaya, M. R. *et al.* *Nature* **447**, 453–457 (2007).
- Sievers, S. A. *et al.* *Nature* **475**, 96–100 (2011).
- Cheng, Y. *Cell* **161**, 450–457 (2015).
- Shi, D., Nannenga, B. L., Iadanza, M. G. & Gonen, T. *eLife* **2**, e01345 (2013).
- Gonen, T. *et al.* *Nature* **438**, 633–638 (2005).
- Nannenga, B. L., Shi, D., Leslie, A. G. & Gonen, T. *Nature Methods* **11**, 927–930 (2014).
- Mitsuoaka, K. *et al.* *J. Mol. Biol.* **286**, 861–882 (1999).

This article was published online on 9 September 2015.

Storms bring ocean nutrients to light

Ships and ocean-observing robots have been used to quantify the amount of nutrients that a storm brings up from the Stygian ocean depths to the sunlit surface — a first step in assessing how storms affect oceanic biomass production.

JAIME PALTER

Seawater absorbs most sunlight within 100 metres of the ocean's surface, where available inorganic nutrients are rapidly assimilated by photosynthesizing plankton. As these cells die or are excreted in a consumer's waste, they sink into the dark ocean interior, where microbial decay returns their nutrients to the dissolved inorganic form. The physical processes that lift the resulting nutrient-rich water from the dark interior to the sunlit upper ocean are responsible for sustaining nearly all marine life. Writing in *Global Biogeochemical Cycles*, Rumyantseva *et al.*¹ quantify how a mid-latitude storm and its after-effects drive upward pulses of nutrients, thereby advancing our understanding of the processes that nourish plankton at the base of the marine food chain.

Illuminating the processes that bring nutrients into the light has been a preoccupation of oceanographers for generations², a pursuit made more urgent by predictions (see refs 3 and 4, for example) that global warming may suppress the upward nutrient supply, with deleterious effects for many plankton. Such predictions are premised on a causal chain linking the warming of near-surface waters to a

strengthening of the density gradient between the surface and interior ocean, which, in turn, increases the amount of energy required to bring subsurface nutrients upwards. However, the future of the physical drivers that provide this energy is rarely considered.

The turbulence caused by stormy seas is thought to be one mechanism for mixing nutrients upwards. This turbulent mixing is rarely documented because of the challenges of conducting ship-based seawater sampling in high winds. Therein lies the novelty of Rumyantseva and co-authors' study, which reports nutrient concentrations, the intensity of turbulent mixing and the velocities of ocean currents before, during and after the passage of a North Atlantic storm. They solved the sampling problem in part by using torpedo-shaped robots called gliders⁵, which continuously measured water-column properties regardless of the weather.

The authors report an approximately tenfold increase in surface nutrient concentrations during the storm, followed by two short-lived bursts of nutrients after the storm. They also observed that the concentration of chlorophyll, the pigment responsible for most photosynthesis, rose by about 50% near the surface

between the last day of the storm and several days later.

The study provides insight into how storms can produce upward nutrient pulses, even after they cease. Wind blowing over the ocean during a storm sets its surface layer in motion, but the interior ocean not far below is relatively unperturbed. This creates a sharp change in the ocean-current velocity (that is, high shear) at the interface between the surface and the interior that leads to instabilities in the circulation and the development of turbulence (Fig. 1a). Such turbulence can mix nutrients into the sunlit region, and, in the case of the observed storm, was responsible for the rapid increase in surface nutrient concentrations.

After storm winds relent, the surface waters oscillate like a pendulum, rocking back and forth while the planet rotates underneath. The circular path traced by the water relative to a fixed point on Earth is called an inertial oscillation. The shear spikes again when the wind aligns with the ocean surface current (Fig. 1b) — which occurred approximately twice a day for a steady wind over the inertial oscillation observed by Rumyantseva and colleagues. The authors report that these periods of alignment coincided with an upward delivery of nutrients at a rate 25 times faster than background. However, given that these bursts of intense turbulent mixing lasted only about an hour each, their total contribution to the nutrient supply was smaller than that delivered during the storm.

The authors suggest that the net effect of non-winter storms might locally contribute up to 30% as much nutrient as is supplied by wintertime convective mixing, the mechanism that sustains the annual algal blooming of the North Atlantic. But measurements of a single event cannot resolve how storms influence biology on the scale of ocean basins. A study⁶

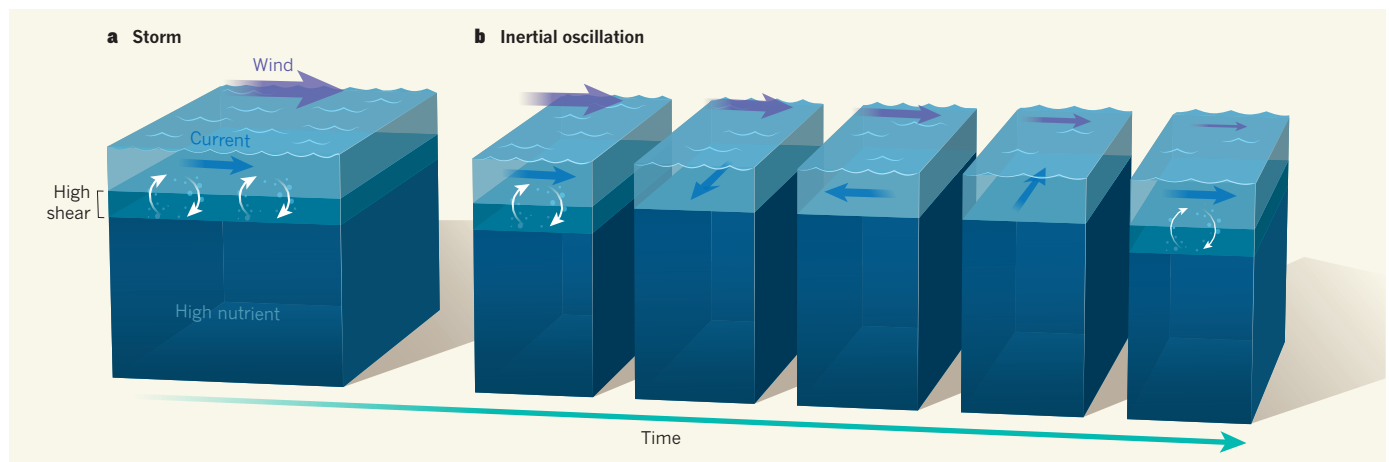


Figure 1 | Uplift of nutrients in the ocean during and after a storm. **a**, During a storm, the current in the upper ocean accelerates in response to the wind passing across it, whereas the interior ocean not far below is relatively unperturbed. This sets up a highly sheared current at the interface between the upper and the interior ocean, which is susceptible to the development of turbulence. Turbulent mixing (white arrows) thus brings nutrients from the interior to the surface. **b**, After the storm, the wind

gradually weakens, and the upper ocean rocks back and forth while the planet rotates underneath. The current therefore traces a circular path relative to a fixed point on Earth — this is known as an inertial oscillation. When the wind and upper-ocean current align, strong turbulent mixing again brings nutrients to the surface. Rumyantseva *et al.*¹ have measured the amount of nutrient brought to the surface during and after a storm in the North Atlantic Ocean.

in the subtropical North Atlantic showed that cyclones drive only a minor increase in the mean amount of chlorophyll during the hurricane season, because their direct influence is felt by a small fraction of the region. This interpretation was recently questioned, however, by another study⁷ that found an association between year-to-year variability of chlorophyll concentration and total cyclone energy during the hurricane season. It therefore remains unclear whether storms are major drivers of such variability on large scales.

Moreover, neither the current paper nor the subtropical studies^{6,7} extended conclusions from chlorophyll concentrations to rates of biological productivity — which is a more relevant metric for understanding the transfer of energy up the food chain, but more challenging to measure. There is thus still much to debate about the magnitude of the large-scale biological response to storms.

Future changes in nutrient supply to the ocean surface will depend on both the energy

required to bring nutrients to the surface and that available from winds and other physical forcing⁸. Climate models widely agree that future warming will strengthen the ocean's vertical density gradient and increase the energy required to mix nutrients upwards⁴. The future of storms is less clear. Climate models generally predict a poleward shift of storm tracks and intensified storms in the Southern Hemisphere, whereas corresponding predictions for the Northern Hemisphere — and particularly the North Atlantic — are highly uncertain⁹.

Attempts to predict the biological influence of storms are further hampered by the inadequate representation of inertial oscillations in climate models, a gap that researchers are now working to fill¹⁰. Such improved numerical representations of how storms influence ocean mixing, and a stronger handle on the large-scale influence of storms on present-day biology, will give us a clearer vision of what the future may hold for marine plankton and the organisms that depend on them. ■

Jaime Palter is at the Graduate School of Oceanography, University of Rhode Island, Narragansett Bay Campus, Narragansett, Rhode Island 02882, USA.
e-mail: jpalter@uri.edu

1. Rumyantseva, A. et al. *Global Biogeochem. Cycles* <http://dx.doi.org/10.1002/2015GB005097> (2015).
2. Redfield, A. C. *Nature* **138**, 1013 (1936).
3. Sarmiento, J. L. et al. *Global Biogeochem. Cycles* **18**, GB3003 (2004).
4. Cabré, A., Marinov, I. & Leung, S. *Clim. Dyn.* **45**, 1253–1280 (2015).
5. Rudnick, D. L. & Cole, S. T. *J. Geophys. Res.* **116**, C08010 (2011).
6. Hanshaw, M. N., Lozier, M. S. & Palter, J. B. *Geophys. Res. Lett.* **35**, L01601 (2008).
7. Foltz, G. R., Balaguru, K. & Leung, L. R. *Geophys. Res. Lett.* **42**, 1158–1164 (2015).
8. Lozier, M. S., Dave, A. C., Palter, J. B., Gerber, L. M. & Barber, R. T. *Geophys. Res. Lett.* **38**, L18609 (2011).
9. Stocker, T. F. et al. in *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* (eds Stocker, T. F. et al.) Tech. Summ. TS5.4.4, TS5.5.3 (Cambridge Univ. Press, 2013).
10. Jochum, M. et al. *J. Clim.* **26**, 2833–2844 (2013).

CARDIAC BIOLOGY

A protein for healing infarcted hearts

Human heart tissue has minimal ability to regenerate following injury. But the protein Fstl1, which is normally expressed in the heart's epicardial region, has now been shown to induce regeneration following heart attack. [SEE ARTICLE P.479](#)

GORDANA VUNJAK-NOVAKOVIC

Healthy mammalian heart tissue has a measurable but limited ability to regenerate¹. Over a normal human lifespan, around 45% of heart-muscle cells (cardiomyocytes) are renewed, with the remaining 55% persisting from birth. This rate is not sufficient to repair the injury caused by myocardial infarction, or heart attack as it is commonly known. Instead, the infarcted area becomes populated by fibroblast cells, which form a non-contractile collagenous scar — a quick fix that progressively decreases the heart's pumping capacity. Any regenerative therapy must thus provide an influx of cells that can properly heal the muscle, either from an external source or from the body itself². A major effort^{3,4} is going into the potential use of immature cardiomyocytes derived from human stem cells for such regeneration. But in this issue, Wei et al.⁵ (page 479) take a different approach, making use of a protein that is present in the epicardial region of a healthy heart but lost following heart infarction.

Previous work documented that the protein follistatin-like 1 (Fstl1) is involved

in the development of many organ systems, by binding to proteins of the transforming growth factor- β (TGF- β) family and inhibiting their functions⁶. Fstl1 is also involved in a spectrum of diseases, from heart attacks to arthritis,

lung fibrosis and cancer, through its activation of multiple signalling pathways and inflammatory and immune responses. Depending on the organ system, Fstl1 can act as a pro-inflammatory molecule⁷ or a cell-protective factor⁸, or can induce immune responses⁵.

Fstl1 is also known as a modulator of cardiac development⁹ and as a marker of heart ischaemia (restricted blood supply), hypertrophy (abnormal enlargement) and end-stage heart failure⁸. The many roles of Fstl1 are seemingly contradictory: it protects cardiomyocytes from apoptotic cell death and hypertrophy by mobilizing signalling through the phosphorylated kinase enzyme AMPK, but it suppresses their differentiation from stem cells by inhibiting signalling of the TGF- β family member BMP. Notably, the presence of Fstl1 in the heart correlates with reduced infarct size

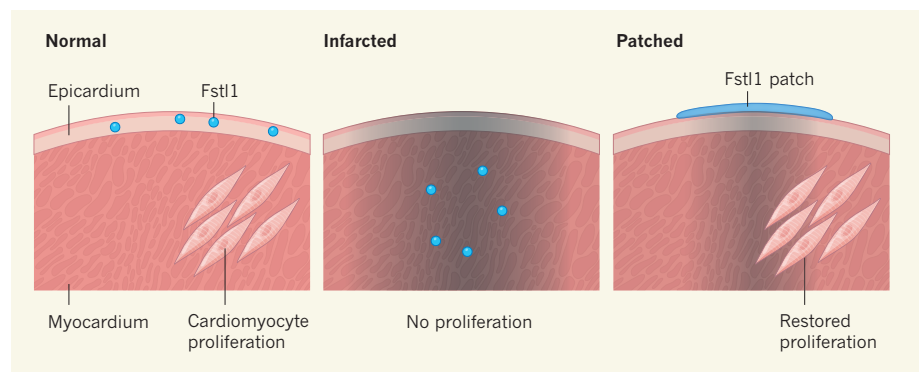


Figure 1 | Patched up to heal. Myocardial infarction (a heart attack) results in a massive loss of heart-muscle cells (cardiomyocytes), which need to be regenerated for the injured site to heal in a way that restores pumping. Wei et al.⁵ show that the protein Fstl1 is typically expressed in the epicardial layer surrounding the myocardium (muscle), and that this protein can induce cardiomyocyte proliferation — but also that this cardiogenic activity is lost when the expression shifts to the myocardium following infarction. However, they show that applying a collagen patch containing epicardial Fstl1 to a mouse heart immediately after infarction can reconstitute the cardiogenic activity of the protein and regenerate the heart muscle.

and with functional recovery, but this effect has been ascribed to enhanced reformation of blood vessels (revascularization) and cell survival, rather than the formation of cardiomyocytes¹⁰.

Wei *et al.* now provide new and counter-intuitive insights into the biological functions of Fstl1. Their study shows that, in the healthy heart, the protein is expressed in the epicardium, the membranous layer surrounding the myocardium, throughout development and in adult life. They also observed that heart infarction causes Fstl1 expression to be transferred from the epicardium to the myocardium, and that this shift impairs the heart's regenerative ability. Remarkably, the study reveals that re-established expression of epicardial Fstl1 can regenerate the injured heart muscle.

The investigators hypothesized that a patch releasing epicardial Fstl1, when placed onto the heart infarct, would serve as a source of Fstl1 and stimulate proliferation of the resident cardiomyocytes (Fig. 1). To test this hypothesis, they loaded collagen patches either with medium containing Fstl1 in which epicardial cells had been cultured, or with human Fstl1 purified from a bacterial protein-expression system, and sutured these to the hearts of mice that had undergone modelled myocardial infarction. Four weeks later, they observed more cardiomyocytes, higher transcription of cardiac marker genes and a greater frequency of calcium pulses (indicative of heart pumping) compared with infarcted hearts without patches. There was also less formation of fibrotic scar tissue and a better revascularization of the area. These findings suggest that such *in situ* manipulation might allow control of the fate of existing cardiomyocytes, to achieve heart regeneration without implanting cells.

This study is an inspiring example of how a developmentally conserved regulatory pathway can be mobilized to induce heart regeneration. Although more work needs to be done to determine the benefits of such an approach in large-animal models (the authors conducted a preliminary study in pigs, but it involved only six animals divided into three groups), the proposed reconstitution of epicardial Fstl1 could lead to entirely new modalities for treating heart infarction. The study also leaves us with questions about the biological phenomena responsible for the observed effects. Fstl1 is still an enigmatic protein with largely unknown properties, but with seemingly huge potential for diagnosing and treating heart disease.

One intriguing question is why infarction-induced myocardial expression of Fstl1, or even experimentally induced overexpression of Fstl1 in the myocardium, cannot induce heart regeneration, but epicardial Fstl1 applied on the patch can. The authors also find this result paradoxical. They suggest that different extents of glycosylation (the number of

carbohydrate molecules — glycans — attached to the protein) that they measured for epicardial and myocardial Fstl1 reflect differences in glycan structure that affect the proteins' function. It remains to be seen whether these differences are cell-of-origin specific, how important glycosylation is for regenerative ability, and what the necessary features would be for a patch that can induce regeneration in the human heart.

Other questions arise from the combined observations that, although myocardial Fstl1 does not induce cardiomyocyte generation, it does protect immature cardiomyocytes, whereas epicardial Fstl1 on a patch enhances cardiomyocyte proliferation, but is not cell-protective. Further investigation is needed to explore whether glycosylation is a key determinant of cardioprotective versus cardiogenic effects, as proposed by Wei and colleagues. Finally, the study suggests that only very immature cardiomyocytes respond to Fstl1. The genetic signatures and the origin of the responsive cells (whether they are resident or recruited) also remain to be determined.

These questions are likely to motivate future studies. Exciting approaches are now emerging at the interface of stem-cell biology and tissue

engineering. High-fidelity models of human heart tissue, combined with findings such as these, could markedly advance quantitative biological research and the clinical translation of discoveries into curative treatments for heart disease. ■

Gordana Vunjak-Novakovic is in the Departments of Biomedical Engineering and Medicine, Columbia University, New York, New York 10032, USA.

e-mail: gv2131@columbia.edu

1. Bergmann, O. *et al.* *Science* **324**, 98–102 (2009).
2. Laflamme, M. A. & Murry, C. E. *Nature* **473**, 326–335 (2011).
3. Chong, J. J. *et al.* *Nature* **510**, 273–277 (2014).
4. Menasché, P. *et al.* *Eur. Heart J.* **36**, 2011–2017 (2015).
5. Wei, K. *et al.* *Nature* **525**, 479–485 (2015).
6. Sylva, M., Moorman, A. F. M. & van den Hoff, M. J. B. *Birth Defects Res. C* **99**, 61–69 (2013).
7. Miyamae, T. *et al.* *J. Immunol.* **177**, 4758–4762 (2006).
8. Ogura, Y. *et al.* *Circulation* **126**, 1728–1738 (2012).
9. Mercola, M., Ruiz-Lozano, P. & Schneider, M. D. *Genes Dev.* **25**, 299–309 (2011).
10. van Wijk, B., Gunst, Q. D., Moorman, A. F. M. & van den Hoff, M. J. B. *PLoS ONE* **7**, e44692 (2012).

This article was published online on 16 September 2015.

NUCLEAR PHYSICS

Neutrons with a twist

Neutrons do not normally have orbital angular momentum. But the demonstration that a beam of neutrons can acquire this property, 23 years after it was shown in photons, offers the promise of improved imaging technologies. SEE LETTER P.504

ROBERT W. BOYD

Neutrons were discovered in 1932 by the physicist James Chadwick, and the particles continue to amaze scientists to this day. It was initially thought that neutrons were elementary particles — that is, that they were not composed of other particles. But we now know that, just like protons, neutrons are comprised of three elementary particles called quarks. Quarks have an intrinsic property known as spin angular momentum (or spin), and they endow the neutron with a spin that has a value of $\frac{1}{2}\hbar$ (where \hbar is the reduced Planck constant). On page 504 of this issue, Clark *et al.*¹ show that a free neutron can have a different kind of angular momentum: orbital angular momentum (OAM).

OAM is a broad concept in modern physics, but is usually associated with the motion of electrons around the atomic nucleus in atoms and molecules. In contrast to spin, OAM is not an intrinsic property of the electron: it can take any value of an integer L multiplied by \hbar ,

whereas the electron's spin has a fixed value of $\frac{1}{2}\hbar$. Electron spin and OAM are analogous to Earth's rotation on its axis and its orbit around the Sun, respectively.

But OAM has also arisen in a different context: in the early 1990s, it was theoretically² and experimentally³ shown that any helically phased light beam can possess OAM. It has since been established that this is true even for a single photon⁴. This is therefore another source of angular momentum for the particle, in addition to its spin (which is associated with the circular polarization of light). It is a crucial property of photons that has found applications in the field of photonics, such as the coding of quantum⁴ and classical⁵ information in individual photons, quantum-entanglement protocols⁶ and the manipulation of small particles by optical forces⁷.

In 2010, electron beams with OAM were also generated, confirming that this property is not limited to light beams⁸. Many advances in the production and use of OAM-carrying electron beams have since been reported (see

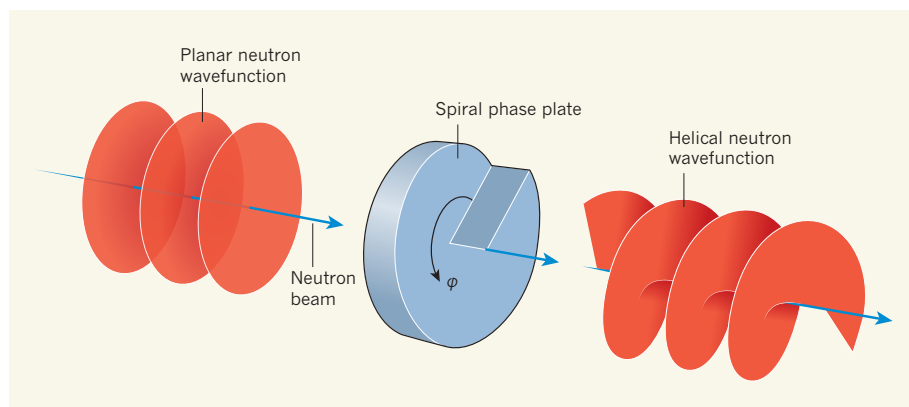


Figure 1 | Orbital angular momentum of neutrons. Clark *et al.*¹ channelled a beam of neutrons through a device known as a spiral phase plate, which modified the neutrons' original, planar wavefunctions and imparted orbital angular momentum to the particles. The wavefunction of the neutrons that emerge from the device has acquired an azimuthal phase distribution of the form $e^{iL\varphi}$ (where i is the imaginary unit, L is any integer and φ is the azimuthal angle of the plate). This phase variation causes the helical structure seen in the emergent wavefunction, which is associated with the acquired orbital angular momentum.

ref. 9 for a review). The fact that photons are not the only particles that can have OAM has opened up possibilities for fundamental studies of electromagnetic interactions and for applications such as improved electron microscopes.

Clark and colleagues' work adds neutrons to the list of particles that can have OAM. The authors generate OAM-carrying neutrons by guiding a beam of the particles through a device known as a spiral phase plate (Fig. 1). The thickness of this device varies uniformly as a function of the plate's azimuthal angle, φ (the angle measured around the circumference of the plate). The wavefunction of a neutron passing through this device acquires a phase shift that is proportional to the plate's local thickness. For appropriate values of the variation of thickness with φ , the wavefunction acquires an azimuthal phase distribution given by $e^{iL\varphi}$, where L is any positive or negative integer and i is the 'imaginary unit' (the square root of -1).

The authors fabricated several plates whose thickness distributions corresponded to various values of L , and thus generated neutron beams carrying OAM of different $L\hbar$ values. Like its spin, a neutron's OAM is a quantum-mechanical attribute. It occurs as a consequence of the helical structure of the particle's 'twisted' wavefunction when it emerges from the plate. To verify that the neutron beam had acquired OAM as it passed through the plate, Clark *et al.* used a technique known as neutron interferometry. In this approach, the neutron wavefunction was split into two paths and a spiral phase plate was placed in one of them. The two paths were subsequently combined coherently to form an output beam whose interference pattern showed the azimuthal phase distribution that the wavefunction had acquired.

Although Clark and colleagues' results are impressive, they represent only the first step in an emerging field of research. For example,

in the present experiment, the neutron beam falling on the spiral phase plate is a statistical mixture of several OAM quantum states. Before applications can be developed, neutrons must be generated that have quantum states with definitive OAM values (eigenstates). In addition, holographic methods have been developed for creating optical^{10,11} and electron¹² OAM states, and these are more precise and versatile than the use of spiral phase plates. It will thus be interesting to explore the use of holographic techniques for neutrons too. The potential use of neutron OAM states for quantum-information studies is another exciting prospect.

Finally, Clark and colleagues' study opens up a further avenue for future work: the use of neutron beams with OAM for imaging. Because neutrons are penetrating particles, they could offer practical advantages compared with optical and electron microscopy in deep-imaging studies of materials. One might therefore conclude that OAM-carrying neutron beams may boldly go where no quantum particle has gone before. ■

Robert W. Boyd is in the Department of Physics, School of Electrical Engineering and Computer Science, and the Max Planck Centre for Extreme and Quantum Photonics, University of Ottawa, Ottawa, Ontario K1N 6N5, Canada, and at The Institute of Optics, University of Rochester, Rochester, New York. e-mail: boydrw@mac.com

1. Clark, C. W. *et al.* *Nature* **525**, 504–506 (2015).
2. Allen, L., Beijersbergen, M. W., Spreeuw, R. J. C. & Woerdman, J. P. *Phys. Rev. A* **45**, 8185–8189 (1992).
3. Beijersbergen, M. W., Allen, L., van der Veen, H. E. L. O. & Woerdman, J. P. *Opt. Commun.* **96**, 123–132 (1993).
4. Mair, A., Vaziri, A., Weihs, G. & Zeilinger, A. *Nature* **412**, 313–316 (2001).
5. Wang, J. *et al.* *Nature Photon.* **6**, 488–496 (2012).
6. Leach, J. *et al.* *Science* **329**, 662–665 (2010).
7. He, H., Friesen, M. E. J., Heckenberg, N. R. &



50 Years Ago

A Biological Retrospect. By Sir Peter Medawar — The title of my presidential address, as you will have discerned, is “A Biological Retrospect”, and on the whole it has not been well received. ‘Why a biological retrospect?’, I have been asked; would it not be more in keeping with the spirit of the occasion if I were to speak of the future of biology rather than of its past? Unfortunately, it is impossible to predict new ideas ... and we are caught in a logical paradox the moment we try to do so. For to predict an idea is to have an idea, and if we have an idea it can no longer be the subject of a prediction. Try completing the sentence ‘I predict that at the next meeting of the British Association someone will propound the following new theory of the relationships of elementary particles, namely...’. If I complete the sentence, the theory will not be new next year; if I fail, then I am not making a prediction.

From *Nature* 25 September 1965

100 Years Ago

We have still ... very much to learn about causes in action; and the mystery of the earth, and of our connection with it, grows upon us as we learn. Can we at all realise the greatest change that ever came upon the globe, the moment when living matter appeared upon its surface ... And here was living matter, a product of the slime, if you will, but of a slime more glorious than the stars. Was this thing, life, a surface-concentration, a specialisation, of something that had previously permeated all matter, but had remained powerless because it was infinitely diffuse? Here you will perceive that the mere geologist is very much beyond his depth.

From *Nature* 23 September 1915

- Rubinstein-Dunlop, H. *Phys. Rev. Lett.* **75**, 826–829 (1995).
8. Uchida, M. & Tonomura, A. *Nature* **464**, 737–739 (2010).
9. Harris, J. *et al. Nature Phys.* **11**, 629–634 (2015).

10. Bazhenov, V. Y., Vasnetsov, M. & Soskin, M. S. *JETP Lett.* **52**, 429–431 (1990).
11. Bolduc, E., Bent, N., Santamato, E., Karimi, E. & Boyd, R. W. *Opt. Lett.* **38**, 3546–3549 (2013).
12. Verbeeck, J., Tian, H. & Schattschneider, P. *Nature* **467**, 301–304 (2010).

EVOLUTIONARY BIOLOGY

Infection elevates diversity

Chromosomal shuffling in parental eggs or sperm can create new characteristics in the next generation. In fruit flies, it seems that mothers with a parasitic infection produce more such recombinant offspring than uninfected mothers.

ANEIL F. AGRAWAL

In most plants and animals, offspring are genetically distinct from their parents. Through the process of recombination, which occurs as sperm and egg cells (gametes) are produced, a parent can mix the two copies of a given chromosome received from its own parents, thus transmitting unique chromosomes to its offspring. Writing in *Science*, Singh *et al.*¹ show that fruit flies produce a higher frequency of offspring with recombinant chromosomes when the mother is infected with a parasite than when it is uninfected. This intriguing observation may be an important piece in the long-standing puzzle of why recombination is so common.

Why should organisms shuffle their genomes through sex and recombination? Natural selection should create an excess of good gene combinations, so ‘undoing’ the work of past selection by rearranging these genotypes seems counterproductive. One possible explanation is that what constitutes a good combination of alleles (gene variants) changes over time. In that case, undoing the work of past selection is beneficial because selection in the future demands something different. This idea requires that selection on gene combinations changes regularly².

Coevolving natural enemies — particularly parasites — might provide just the right type of selection pressures for this scenario. This is the basis for the ‘Red Queen’ hypothesis, which proposes that sexual reproduction and recombination are favoured because they help hosts to adapt to the ever-shifting selection imposed on their gene combinations by the parasites^{3,4}. However, even rapidly evolving parasites do not always induce selection for recombination; there are times in the coevolutionary cycle when hosts are well adapted and non-recombinant offspring will be more resistant to infection than recombinant ones^{5,6}. Intuitively, it might seem that the ideal solution is to

increase recombination when infected because being infected indicates that your current gene combination is not working.

To test this, Singh *et al.* performed hundreds of test crosses using female *Drosophila melanogaster* fruit flies that carried mutations at each of two genes on one chromosome, but that had normal versions of the genes on their other copy of the chromosome. The presence of either mutation leads to visible physical characteristics that allow determination of whether one or both mutations are present in their offspring — because the genes are in close physical proximity, the normal or mutated versions will be inherited together unless there has been recombination (Fig. 1).

The females were injected with one of two bacterial pathogens (*Serratia marcescens* or *Providencia rettgeri*) or given a sham injection. By examining tens of thousands of the flies’ progeny, the authors found that infected mothers produced a higher fraction of recombinant offspring than non-infected mothers. This effect was seen in four fly strains. Infection with a parasitoid wasp (*Leptopilina clavipes*) also induced an increase in recombinant progeny. Unlike the bacterial experiments, in which reproductive adults were infected, the parasitoid wasp infects fruit-fly larvae and the parasites must be killed for the larva to survive to adulthood. Thus, in this situation, the infection is cleared long before meiosis (the cell division necessary to produce gametes and during which recombination occurs).

An increase in the observed frequency of recombinant progeny from infected mothers could be due to an increase in the recombination rate or to transmission distortion (for example, if recombinant chromosomes are more likely than non-recombinant chromosomes to end up in successful gametes). To tease these possibilities apart, Singh *et al.* made use of the fact that exchange of chromosomal material (crossover events) occurs 4–5 days before eggs are laid. In their bacterial-infection experiments, the authors found an increase in recombinant progeny even in the first 4 days after the mothers were infected. This rapid response points to transmission distortion. A remaining challenge will be to understand how this distortion occurs. Are recombinant chromosomes less likely than non-recombinant ones to end up in polar bodies, the small

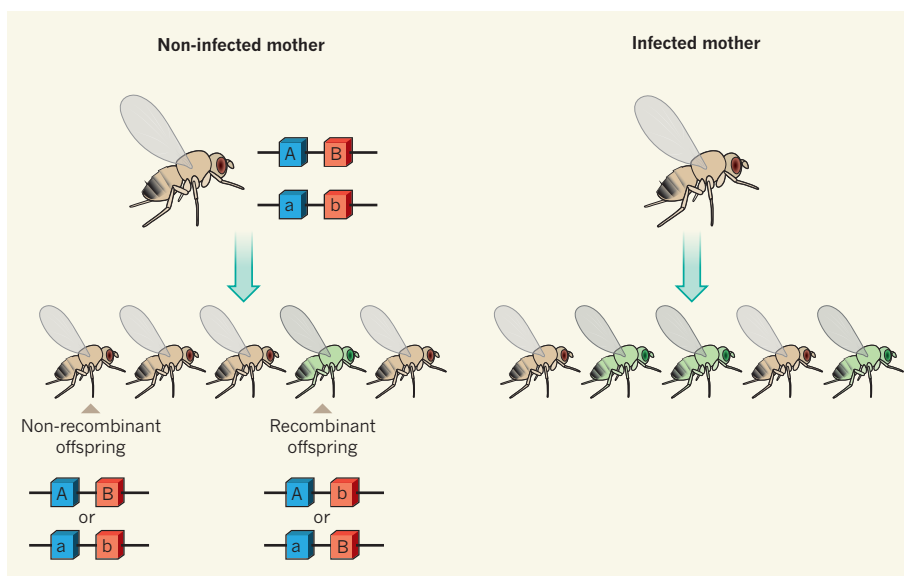


Figure 1 | Frequency of recombinant offspring altered by infection. Diploid organisms, such as fruit flies and humans, have two copies of each chromosome, which can vary in DNA sequence (represented by A versus a and B versus b) in every cell except gametes (sperm and egg cells). Gametes contain only one copy of each chromosome, such that fertilization results in two copies again in the offspring. The sequence in the offspring can be the same as the parental chromosome, or an exchange of genetic material between the two chromosomes during gamete production — recombination — can result in different sequences. Singh *et al.*¹ show that fruit-fly mothers that are infected with parasites produce more recombinant offspring than uninfected mothers.

cells that are formed during meiosis but do not transmit genes to future generations? Are gametes bearing recombinant chromosomes more viable, or do they somehow outcompete gametes that have non-recombinant chromosomes? What mechanism might mediate such a bias? Furthermore, Singh and colleagues' study focuses on a single genomic region — it will be of interest to assess whether similar responses occur elsewhere in the genome, and if not, why this region is particularly responsive.

Previous work has demonstrated that pathogens increase recombination in plants during meiotic and somatic (non-gamete) division^{7–9}. That the few existing examples of this phenomenon span plants and animals suggests that pathogen-induced increases in the recombinant fraction could be widespread, although perhaps achieved through different means, for example transmission distortion in flies but higher recombination in plants. If so, does this intriguing connection between pathogens and natural variation in recombination constitute convincing evidence to support the Red Queen hypothesis?

Changes in the proportion of recombinant

offspring in flies and other organisms have been reported in response to various types of environmental stress (such as temperature, nutrition and social stress; reviewed in ref. 10), although rarely with the rigour of Singh and colleagues' work. Is selection by parasites a driver of the evolution of plasticity in the recombinant fraction, perhaps one that spills over to other types of stress? Or is the observed response to pathogens a by-product of whatever causes plasticity in response to these other stresses? A crucial first step towards answering these questions would be to obtain evidence — so far lacking — that recombinant offspring are less likely to become infected than non-recombinant offspring.

Although plasticity in the recombinant fraction has been known for around 100 years, it is still poorly studied. We have only the crudest picture of what conditions alter the recombinant fraction, by how much and in which genomic regions. Moreover, theoretical models¹⁰ suggest that the evolution of recombination plasticity is not easily explained for 'normal' stresses in diploid organisms (those that have two copies of each chromosome,

including flies). Even the seemingly intuitive Red Queen interpretation of Singh and colleagues' results is questionable because offspring will always receive half the alleles carried by their mother, regardless of whether they are recombinant or not. Although studies such as this shed light on variation in recombination, there is a long way to go in terms of fully describing this variation and understanding it from both a mechanistic and an evolutionary perspective. ■

Aniel F. Agrawal is in the Department of Ecology & Evolutionary Biology, University of Toronto, Toronto, Ontario M5S 3B2, Canada. e-mail: a.agrawal@utoronto.ca

1. Singh, N. D. *et al. Science* **349**, 747–750 (2015).
2. Bell, G. & Maynard Smith, J. *Nature* **328**, 66–68 (1987).
3. Jaenike, J. *Evol. Theory* **3**, 191–194 (1978).
4. Hamilton, W. D. *Oikos* **35**, 282–290 (1980).
5. Peters, A. D. & Lively, C. M. *J. Evol. Biol.* **20**, 1206–1217 (2007).
6. Agrawal, A. F. *Evolution* **63**, 2131–2141 (2009).
7. Kovalchuk, I. *et al. Nature* **423**, 760–762 (2003).
8. Lucht, J. M. *et al. Nature Genet.* **30**, 311–314 (2002).
9. Andronic, L. *Can. J. Plant Sci.* **92**, 1083–1091 (2012).
10. Agrawal, A. F., Hadany, L. & Otto, S. P. *Genetics* **171**, 803–812 (2005).

COMPUTATIONAL ASTROPHYSICS

Monstrous galaxies unmasked

The enigma of how the most luminous galaxies arise is closer to being solved. New simulations show that these are long-lived massive galaxies powered by prodigious gas infall and the recycling of supernova-driven outflows. SEE LETTER P.496

ROMEEL DAVÉ

Three billion years after the Big Bang, the Universe was a different place from today. During that epoch, known as cosmic noon, the average star-formation rate across the cosmos was 100 times higher than it is at present, and individual galaxies were growing commensurately rapidly. This was illustrated by the surprising discovery¹, more than a decade ago, of galaxies whose star-formation rates during that era were 1,000 times the Milky Way's current output — no such galaxies are seen in the present-day Universe. On page 496 of this issue, Narayanan *et al.*² present numerical simulations that offer unprecedented clarity in understanding the origins of such deep-space monsters.

These galaxies have extreme properties and are the most luminous in the Universe. However, despite their enormous total energy output, they are faint at optical wavelengths: most of the radiation emitted by their stars is absorbed by a 'mask' of interstellar

dust and re-emitted at longer wavelengths. Consequently, they remained undiscovered until the advent of surveys at submillimetre and radio wavelengths¹. The very existence of these submillimetre galaxies (SMGs) presented a challenge to models of galaxy formation in a cosmological framework, and has since sparked a vigorous debate in the field of galaxy-formation theory.

Two schools of thought emerged, centred around the 'merger-starburst' and the 'smooth-accretion' hypotheses, respectively^{3–6}. The former proposes that a given SMG is the product of a collision between two gas-rich disk galaxies — this process drives a short-lived (about 10⁸ years) but spectacular burst of star formation during the galaxies' coalescence. The latter argues that SMGs represent the most massive members of the entire galaxy population, being long-lived phenomena that are continuously fed by gas accretion over periods of about 10⁹ years.

The merger-starburst hypothesis stems from a scaled-up analogy of the observation

that the most luminous galaxies in the present-day Universe are almost always involved in spectacular collisions⁷. The smooth-accretion hypothesis is founded on the prediction³ that, at early cosmic epochs, galaxies were accreting gas at extremely high rates — they could thus potentially sustain their excessive star-formation activity.

Neither scenario has been successful in fully replicating the observed properties of SMGs. Researchers have been unable to reproduce the number of systems required to match the observations under the merger-starburst hypothesis⁶, because collisions between sufficiently large galaxies were rare at those early times. An influential study⁴ has argued that collisions between more-numerous low-mass galaxies could lead to the formation of SMGs, but only under the assumption that the star-formation process stimulated by the mergers was heavily weighted towards the production of massive stars. However, this assumption was subsequently disfavoured by observational results⁸.

Similarly, the smooth-accretion scenario has been tested in cosmological simulations that generated SMGs matching those observed, but they did not reproduce the high luminosities of these systems^{5,6}. Given that SMGs are thought to be the progenitors of the well-studied elliptical galaxies found in the present-day Universe, the inability to fit these sources straightforwardly into a cosmological galaxy-formation context has been worrying.

In the current study, Narayanan *et al.* present a hydrodynamic simulation in a cosmological framework that yields the first SMG with a luminosity that is a good match to the

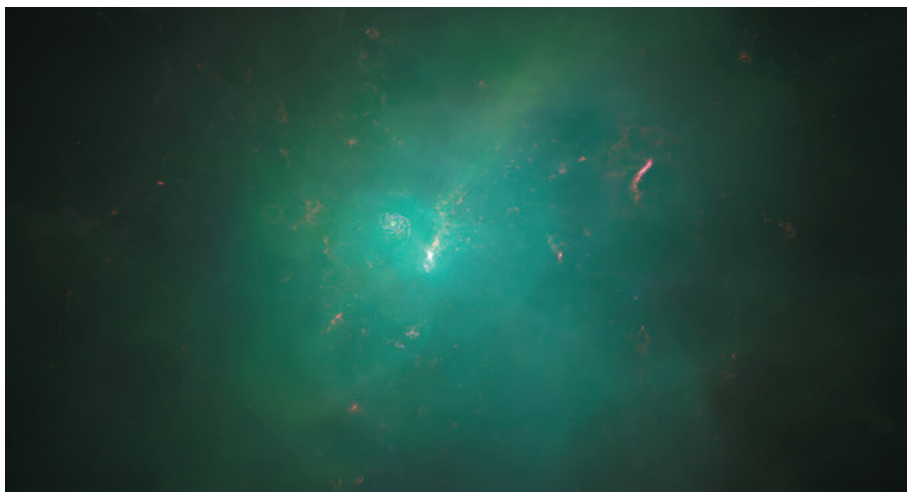


Figure 1 | Simulation of a submillimetre galaxy (SMG). Narayanan *et al.*² simulate how the Universe's most luminous galaxies, which look extremely bright in the submillimetre part of the spectrum, may have formed when the Universe was 3 billion years old. This snapshot, taken from a supercomputer simulation, depicts the distribution of gas and light in a small region of the field: it contains a bright central galaxy (white) that is accreting gas along a filamentary structure (pink), a large spiral galaxy (left of centre), and numerous smaller galaxies that contribute to the total luminosity of the SMG. Ambient gas (blue-green), much of which was expelled by the galaxies at earlier epochs, gravitates towards the centre of the proto-SMG. This fuels the prodigious star-formation activity of the system, which is unlike anything seen in the present-day Universe.

observations, and that can, by extrapolation, reproduce the numbers of SMGs that are observed. This simulation achieves levels of realism that previous models lacked (Fig. 1), by using a 'zoom' technique, in which the authors resimulate a selected region at much higher spatial resolution than the whole simulated volume. They thus obtain an accurate representation of the galaxy-assembly process on subgalactic scales, while producing models that retain the full cosmological context. Furthermore, the high resolution enables the authors to develop a fully self-consistent description of galactic outflows driven by the supernova explosions of massive stars that eject copious quantities of gas into intergalactic space — a key aspect in regulating early galaxy growth. The authors use a newly developed code for radiative transfer to accurately predict the energy output of the SMG at submillimetre wavelengths. These crucial improvements in modelling lead to a deeper understanding of SMGs.

So what do the authors find? The key outcome of their simulation is a 'long-lived' SMG that can sustain star-formation rates of 500–1,000 solar masses per year for about 10^9 years, and which has a submillimetre luminosity that matches typical observations. Such a demonstrably rapid growth in stellar mass results in SMGs that are among the most massive objects in the Universe at cosmic noon. This realization has two corollaries: first, the strong gravitational attraction of the sources causes numerous other galaxies to cluster around them, thus adding non-trivially to the system's total submillimetre luminosity. Second, galactic outflows cannot easily escape the intense gravitational pull of an SMG, but instead rain

back down on the galaxy. This process provides additional fuel that enhances the star-formation rate. In a nutshell, the authors find that SMGs plausibly arise from a 'perfect storm' of high rates of gravitationally driven gas accretion, the recycling of previously ejected material, and contributions to the systems' submillimetre luminosity from nearby galaxies that cannot be well resolved observationally.

Narayanan and colleagues' results favour the smooth-accretion scenario over the merger-starburst hypothesis for the formation of SMGs, and provide key insights into

this decade-old debate. This does not mean that galaxy mergers cannot create SMGs; they probably do. But the current work suggests that they are a minority of cases. What is particularly encouraging is that the authors did not tune the simulations so as to reproduce SMGs: rather, they simply used a state-of-the-art galaxy-formation model and ran it at the highest currently feasible numerical resolution — and a plausible SMG emerged.

Although this study is important, it is unlikely to be the final word on the formation of SMGs. Predicting the cosmic abundance of such galaxies from the simulation of a single object still requires uncertain extrapolations, and the authors' prescription for generating galactic outflows is not unique; other prescriptions may yield different results. These limitations notwithstanding, Narayanan *et al.* have presented the first impressively viable model of SMG formation, allowing us a tantalizing glimpse behind the mask of these behemoths of deep space. ■

Romeel Davé is in the Department of Physics, University of the Western Cape, Bellville 7535, Cape Town, South Africa, and at the South African Astronomical Observatory and the African Institute for Mathematical Sciences. e-mail: romeeld@gmail.com

1. Blain, A. W., Smail, I., Ivison, R. J., Kneib, J.-P. & Frayer, D. T. *Phys. Rep.* **369**, 111–176 (2002).
2. Narayanan, D. *et al. Nature* **525**, 496–499 (2015).
3. Dekel, A. *et al. Nature* **457**, 451–454 (2009).
4. Baugh, C. M. *et al. Mon. Not. R. Astron. Soc.* **356**, 1191–1200 (2005).
5. Finlator, K., Davé, R., Papovich, C. & Hernquist, L. *Astrophys. J.* **639**, 672–694 (2006).
6. Davé, R. *et al. Mon. Not. R. Astron. Soc.* **404**, 1355–1368 (2010).
7. Sanders, D. B. & Mirabel, I. F. *Annu. Rev. Astron. Astrophys.* **34**, 749–792 (1996).
8. Tacconi, L. J. *et al. Astrophys. J.* **680**, 246–262 (2008).

EPIGENETICS

The karma of oil palms

Despite their clonal origin, some oil palm trees develop fruits that give almost no oil. It emerges that the number of methyl groups attached to a DNA region called Karma determines which plants are defective. SEE LETTER P.533

JERZY PASZKOWSKI

Vegetative propagation is a form of asexual reproduction that is routinely used for the commercial mass production of garden plants and trees, because it enables rapid multiplication of highly performing, genetically identical individuals. For certain species, vegetative propagation is hugely demanding and requires technically sophisticated aseptic cultures that produce large numbers of cloned embryos which can

develop into plantlets. But a proportion of plants propagated in this way display developmental abnormalities caused either by genetic aberrations or by epigenetic changes, which stably alter the expression of genes without affecting the underlying DNA sequence¹. On page 533 of this issue, Ong-Abdullah *et al.*² describe a culture-induced epigenetic defect in oil palms caused by specific losses in the number of methyl groups attached to a particular region of their DNA.

High-yielding varieties of oil palm, which

are grown in East Asia, are propagated through tissue-culture techniques that regenerate plants from specific parts of the leaf. These clonal, genetically identical trees are supplied to plantations. Some, however, known as ‘mantled’ palms, develop abnormal flowers and yield much less oil. But young palms need several years of intensive care before they start to fruit, and it is only then that mantling can be detected.

Because of the widespread use of palm oil for a variety of household, food and cosmetic products, the mantled defect is a serious economic problem. Therefore, the way in which this trait is inherited has been well studied. Mantled palms do not follow the Mendelian rules of inheritance, suggesting that the defect is a result of epigenetic changes to gene expression rather than a straightforward gene mutation³. Unfortunately, epigenetic changes are more difficult to pinpoint than genetic lesions. Nonetheless, some promising clues to the cause of mantling have been gathered.

The flowers of mantled trees resemble those found in a mutated form of the model plant *Arabidopsis*. The gene that is defective in the mutant *Arabidopsis* encodes a factor essential for the formation of flower organs, and the corresponding gene in the oil palm has been identified as *EgDEF1* (ref. 4). In mantled flowers, the expression of *EgDEF1* is reduced⁵. Such alterations in gene expression can be caused by DNA methylation, an epigenetic modification in which methyl groups become attached to DNA. Because the vegetative propagation of palms by tissue culture causes a general reduction in levels of DNA methylation³, the suppression of *EgDEF1* expression is likely to be indirectly regulated by loss of DNA methylation. But the question of how this occurs remains.

Genes are surrounded by, and sometimes even peppered with, elements derived from ancient viruses that have invaded genomes over evolutionary time. Most of these elements became inactive and, accordingly, their DNA is heavily methylated. But the geneticist Barbara McClintock, who discovered these elements in maize (corn) more than 60 years ago⁶, found that they are sometimes expressed, and can even move to new chromosomal locations, where they can interfere with gene expression. McClintock believed that these transposable elements perform key regulatory functions within the host genome⁷. *EgDEF1* contains two transposable elements that have been studied previously, but the levels of DNA methylation and the activity of these elements are not linked to mantling⁵. So the molecular mechanism causing mantling has remained a mystery — until now.

Ong-Abdullah *et al.* performed an unbiased genome-wide search for alterations in DNA methylation that were tightly linked to the mantled trait. Key to their experimental design was the analysis of four groups of

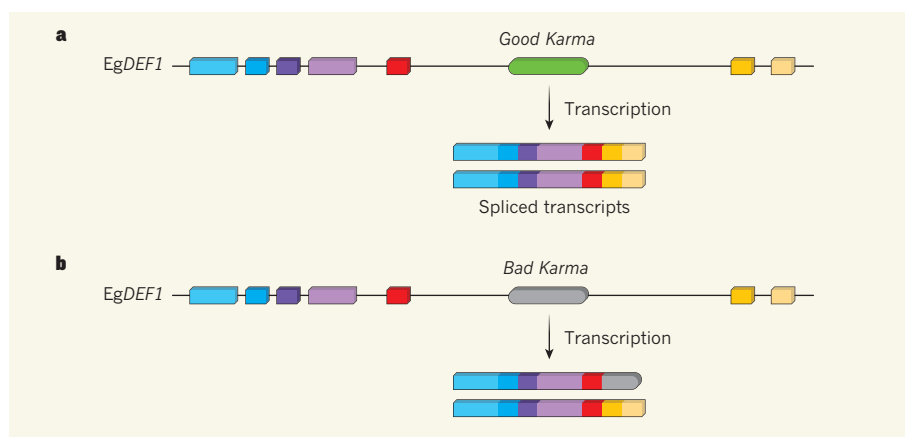


Figure 1 | A mechanism for mantling. Propagation of oil palms can produce defective ‘mantled’ plants. Ong-Abdullah *et al.*² report that mantling is mediated by the number of methyl groups attached to the DNA of a transposable element called *Karma* within the gene *EgDEF1*. **a**, In healthy palms, *Karma* is heavily methylated (*Good Karma*). The element is inactive, and so full-length transcripts are produced — this includes every protein-coding sequence (coloured regions) of *EgDEF1* and omits every non-coding sequence (black connecting lines), including *Good Karma*. **b**, A reduction in methylation leads to *Bad Karma*. This causes alternative splicing of *EgDEF1* RNA to produce an extra transcript that ends at *Karma*, and reduces the number of full-length transcripts produced. The aberrant transcript might be translated to give a truncated protein, which may be responsible for mantling.

palms that differed in their genetic make-up — this reduced the number of false positives and increased the accuracy of the authors’ approach. These genome-wide analyses yet again pointed to the *EgDEF1* gene, but this time methylation changes associated with mantling were detected in a fragment of the gene that had previously been overlooked. Ong-Abdullah and colleagues discovered that this fragment, which lies in a long non-protein-coding region, contains a third transposable element, called *Karma*.

The researchers demonstrated that *Karma*’s DNA remains methylated in healthy plants (a situation they dubbed *Good Karma*) but that methylation is reduced in mantled palms (*Bad Karma*). Remarkably, *Karma* encodes a ‘splice acceptor’ site — a sequence that directs splicing in the gene’s RNA transcript. The authors report that the *Karma* splice site is used only when its methylation is reduced. Although the mechanisms underlying this specificity are not known, Ong-Abdullah *et al.* did show that mantled flowers produce an alternatively spliced *EgDEF1* transcript that accumulates during flower development, and that may encode a truncated *EgDEF1* protein (Fig. 1).

Taken together, these data show that the loss of DNA methylation at *Karma*, and the subsequent *Karma*-mediated alternative splicing, are linked to the mantled trait. It is still not clear whether mantling is triggered by the production of truncated protein, by accumulation of an aberrant transcript spliced at *Karma*, by a reduction in *EgDEF1* transcript levels or by a combination of all these factors. But whatever the exact role of *Karma*, we now know that it can be *Good Karma*, methylated and harmless, or *Bad Karma*, demethylated

and associated with mantling.

Ong-Abdullah and colleagues’ finding is likely to provide a way to detect economically worthless palms much earlier than was previously possible, enabling their timely replacement in plantations. This would be not only of obvious economic importance, but also of relevance to the environment. Oil-palm plantations take over the precious space of tropical forests and any increase in their productivity will contribute to the sustainability of palm-oil production.

The results also have other key implications. For instance, they show that well-planned and performed genome-wide methylation mapping can pinpoint precise spots in the genome of a non-model organism that are responsible for a trait of interest. This paves the way for similar studies that could shed light on the issue of ‘missing’ heritability⁸. Moreover, this approach might lead to more examples of key regulatory roles for transposable elements, lending support to the predictions made by McClintock decades ago⁶. ■

Jerzy Paszkowski is at the Sainsbury Laboratory, University of Cambridge, Cambridge CB2 1LR, UK.
e-mail: jerzy.paszkowski@slcu.cam.ac.uk

1. Stroud, H. *et al.* *eLife* **2**, e00354 (2013).
2. Ong-Abdullah, M. *et al.* *Nature* **525**, 533–537 (2015).
3. Jalignot, E. *et al.* *Ann. Bot.* **108**, 1453–1462 (2011).
4. Adam, H. *et al.* *J. Exp. Bot.* **58**, 1245–1259 (2007).
5. Jalignot, E. *et al.* *PLoS ONE* **9**, e91896 (2014).
6. Profiles in Science: The Barbara McClintock Papers, ‘Controlling elements: Cold Spring Harbor, 1942–1967’, available online at <http://profiles.nlm.nih.gov/ps/retrieve/Narrative/LL/p-nid/49>
7. McClintock, B. *Science* **226**, 792–801 (1984).
8. Manolio, T. A. *et al.* *Nature* **461**, 747–753 (2009).

This article was published online on 9 September 2015.

Hallmarks of pluripotency

Alejandro De Los Angeles^{1,2,3}, Francesco Ferrari⁴, Ruibin Xi^{4,5}, Yuko Fujiwara^{1,2,3}, Nissim Benvenisty⁶, Hongkui Deng⁷, Konrad Hochedlinger^{2,3,4,8}, Rudolf Jaenisch⁹, Soohyun Lee⁴, Harry G. Leitch¹⁰, M. William Lensch^{1,2,3}, Ernesto Lujan¹¹, Duanqing Pei¹², Janet Rossant¹³, Marius Wernig¹¹, Peter J. Park⁴ & George Q. Daley^{1,2,3}

Stem cells self-renew and generate specialized progeny through differentiation, but vary in the range of cells and tissues they generate, a property called developmental potency. Pluripotent stem cells produce all cells of an organism, while multipotent or unipotent stem cells regenerate only specific lineages or tissues. Defining stem-cell potency relies upon functional assays and diagnostic transcriptional, epigenetic and metabolic states. Here we describe functional and molecular hallmarks of pluripotent stem cells, propose a checklist for their evaluation, and illustrate how forensic genomics can validate their provenance.

Stem cells, defined by dual hallmark features of self-renewal and differentiation potential, can be derived from embryonic and postnatal animal tissues and are classified according to their developmental potency (Fig. 1). The zygote and blastomeres are totipotent¹, denoting potential to give rise to all embryonic and extra-embryonic tissues, but their developmental potential has not been captured *in vitro*. Mouse embryonic stem cells exemplify a quintessential pluripotent stem (PS) cell that can form all tissues of the body, but provides only limited contributions to the extra-embryonic membranes or placenta. As described in greater detail below, PS cells manifest distinct functional properties depending upon the conditions under which they are derived and cultured. Multipotent stem cells, such as the paradigmatic haematopoietic stem cell, are restricted to generating the mature cell types of their tissue of origin, but under normal physiologic circumstances will not differentiate into unrelated lineages. Unipotent stem cells, such as spermatogonial stem cells (SSCs), share the capacity for self-renewal yet exhibit limited developmental potential, giving rise to only a single cell type, such as sperm.

Human PS cells correspond to a stable state allowing propagation of immortal pluripotent cells that can generate any cell within the body. Nuclear reprogramming, via somatic cell nuclear transfer and transcription factor transduction, demonstrates that the specialized state of a somatic cell can be reversed to a totipotent or pluripotent state, respectively^{2,3}. The generation of induced pluripotent stem (iPS) cells from somatic cells via transcription factor expression constitutes a facile route to generate patient-specific PS cells, and has opened new paths to model diseases and new prospects for regenerative medicine. Given their versatility for medical applications, PS cells command considerable attention; therefore, defining the hallmarks of pluripotency has practical as well as fundamental value to biomedical research.

In this technical review, we describe the hallmark characteristics of PS cells, propose a checklist of assays for assessing the function and molecular state of pluripotency, and outline forensic genomic approaches to validate the provenance of reprogrammed cell lines.

Defining pluripotent stem cells

PS cells are self-renewing cells with the capacity to form representative tissues of all three germ layers of the developing embryo—ectoderm,

mesoderm and endoderm, as well as the germ lineage, but typically provide little or no contribution to the trophoblast layers of placenta. PS cells can be derived from numerous sources (Table 1). The first PS cells cultured *in vitro* were derived from teratocarcinomas, a tumour of germ cell origin⁴. Later, derivation of PS cells from the murine blastocysts proved that pluripotent cells could be propagated as immortalized, non-transformed cell lines^{5,6}. PS cells have also been derived from non-human primate and human embryos^{7,8}, and from various stages of development, including the post-implantation epiblast and the germ line^{9–14}. Finally, somatic cells can be reprogrammed to pluripotency by ectopic expression of select sets of transcription factors³.

PS cells manifest distinct properties depending on derivation and maintenance conditions. PS cells established from pre-implantation embryos are known as ES cells, whereas those generated from slightly later embryonic epiblast stages are called epiblast stem cells (EpiSCs)^{9,10}. Their distinct culture requirements, gene expression programs and epigenetic features may reflect the dynamic development of pluripotency in the embryo. The terms ‘naive’ and ‘primed’ were introduced to describe early and late phases of epiblast ontogeny and respective ES cell and EpiSC derivatives¹⁵. PS cells from various sources have been classified accordingly (Table 1). Conventional human PS cells exhibit molecular attributes similar to EpiSCs and are classified as ‘primed’. Evaluation of naive pluripotency in humans by formation of human chimaeras is restricted on ethical grounds in many jurisdictions, but as conventional non-human primate ES cells fail to chimaerize pre-implantation embryos, traditional human ES cells are also probably primed by this criterion¹⁶.

Molecular hallmarks of pluripotency

PS cells are characterized by molecular mechanisms that sustain self-renewal and suppress differentiation while maintaining key differentiation genes in a quiescent yet ‘poised’ state reflective of their incipient developmental potential.

A select set of core transcription factors in combination governs and thereby defines pluripotency: OCT4 (also known as POU5F1), SOX2 and NANOG (collectively, OSN). OCT4 and NANOG are designated as

¹Stem Cell Transplantation Program, Division of Pediatric Hematology Oncology, Children’s Hospital Boston, and Dana-Farber Cancer Institute; Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, Massachusetts 02115, USA. ²Harvard Stem Cell Institute, Cambridge, Massachusetts 02138, USA. ³Howard Hughes Medical Institute, Boston, Massachusetts 02115, USA. ⁴Department of Biomedical Informatics, Harvard Medical School, Boston, Massachusetts 02115, USA. ⁵School of Mathematical Sciences and Center for Statistical Science, Peking University, Beijing 100871, China. ⁶Stem Cell Unit, Department of Genetics, Institute of Life Sciences, The Hebrew University, Jerusalem 91904, Israel. ⁷College of Life Sciences and Peking-Tsinghua Center for Life Sciences, Peking University, Beijing 100871, China. ⁸Massachusetts General Hospital Cancer Center and Center for Regenerative Medicine, Boston, Massachusetts 02114, USA. ⁹Whitehead Institute for Biomedical Research, Cambridge, Massachusetts 02142, USA. ¹⁰Medical Research Council Clinical Sciences Centre, Imperial College London, London W12 0NN, United Kingdom. ¹¹Institute for Stem Cell Biology and Regenerative Medicine, Stanford University, Stanford, California 94305, USA. ¹²South China Institute for Stem Cell Biology and Regenerative Medicine, Guangzhou Institutes of Biomedicine and Health, Chinese Academy of Sciences, Guangzhou 510530, China. ¹³The Hospital for Sick Children Research Institute, Toronto, Ontario ON M5G 0A4, Canada.

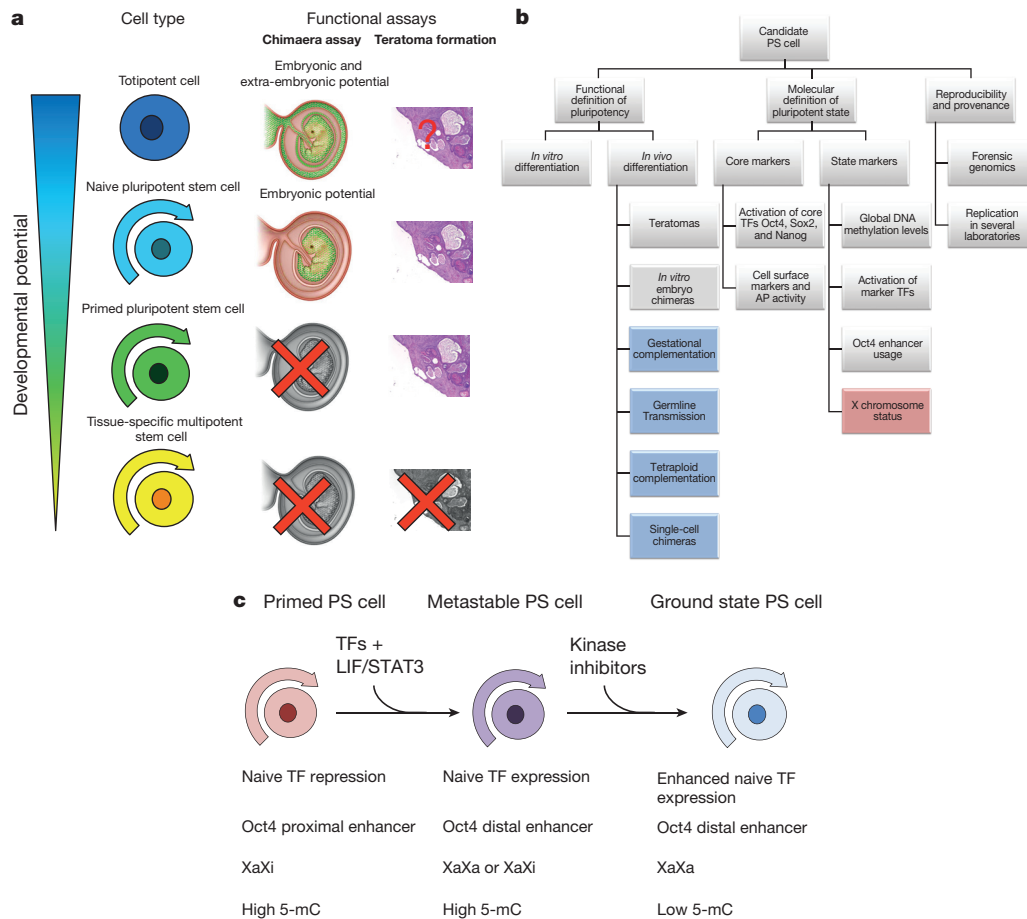


Figure 1 | Stem-cell potency. **a**, Two cardinal assays for assessing PS-cell potency are blastocyst chimaerism and teratoma formation. Performance in these assays allows classification of totipotent, naive pluripotent, primed pluripotent, and multipotent developmental potentials. Totipotency is defined by the capacity to develop and form all tissues of the organism, including extra-embryonic tissues. Naive PS cells are distinguished by the capacity to form a teratoma and a chimaeric animal following introduction into pre-implantation embryos, whereas primed PS cells form teratomas but do not efficiently form chimaeras following introduction into pre-implantation embryos. Tissue-specific multipotent stem cells form cell types related to their tissue-of-origin, but do not form teratomas or chimaeras. Primed EpiSCs do not efficiently form chimaeras when introduced into blastocysts, but can contribute to non-viable post-implantation chimaeras. Therefore, EpiSCs also exhibit pluripotency when introduced into post-implantation embryos. A strict criterion for potency is the demonstration that a single cell can differentiate into the different cell types via single-cell transplantation or by genetically labelling test cells and demonstrating that the daughters of a single cell contribute to different lineages. For human PS cells, teratoma formation remains the gold standard functional assay. Although single-cell-derived teratomas have not been directly generated from diploid human PS cells, clonal-cell-line-derived teratomas provide indirect evidence for the developmental potential of human PS cells at a single-cell level. **b**, Checklist for assessing the function and state of candidate PS cells. Validating the pluripotency of novel PS cells involves assessment of 'function' by measuring self-renewal capacity and developmental potential, and validating pluripotency as a 'state' by measuring the activation of core pluripotency transcription factors (TFs) OCT4, SOX2 and NANOG, and characterization of state markers, such as marker transcription factors and

DNA methylation levels. For example, human ground state PS cells are anticipated to exhibit global DNA hypomethylation and reactivation of transcription factors expressed during pre-implantation development. For novel claims of PS cells, when possible, forensic-genomics-based approaches and independent reproduction in an independent laboratory should validate the provenance and reproducibility of pluripotent phenomena. The blue boxes indicate *in vivo* differentiation assays that should not be assessed in human cells; the red box indicates the uncertain relevance of X chromosome reactivation as a criterion for human ground state PS cells owing to the unresolved interpretation of X chromosome status in human naive pluripotency. AP activity, alkaline phosphatase activity. **c**, Resetting to ground state pluripotency. Primed PS cells exhibit high levels of DNA methylation, cannot chimaerize pre-implantation blastocysts, and female primed PS cells exhibit post-X-chromosome-inactivation status. Xa, active X chromosome; Xi, inactivated X chromosome. To overcome the differentiation barrier between naive and primed PS cells, transcription factors (TFs) are introduced into primed PS cells to initiate resetting. Transcription-factor-induced PS cells or metastable PS cells cultivated in ground state culture conditions will be reset to ground state pluripotency, demarcated by homogeneous expression of naive transcription factors and global DNA hypomethylation (low 5-mC) reminiscent of pre-implantation embryo cells. Globally hypomethylated genomes in ground state mouse ES cells resemble pre-implantation blastocysts, whereas serum-cultivated mouse ES cells and primed EpiSCs possess a hypermethylated genome reminiscent of post-implantation epiblasts and somatic cells. The methylation state of altered human PS cells is undefined, but reset cells generated by the Smith laboratory exhibit DNA methylation level changes closer to ground state mouse ES cells⁵⁵.

core transcription factors based on their specific expression pattern in PS cells and early embryos, and genetic screens identifying their essential role in establishing pluripotency in mice and humans^{3,17–26}. OCT4 functions as a heterodimer with SOX2, placing SOX2 among the core regulators²². The generation of mouse and human iPS cells by ectopic expression of OCT4 and SOX2 highlights the pre-eminent role of OCT4/SOX2 in establishing pluripotency. Although NANOG is not required for mouse PS-cell

maintenance²⁵, and is expressed at low or absent levels in mouse EpiSCs, it stabilizes PS cells, is necessary for *in vivo* pluripotency to develop in the inner cell mass (ICM)²⁶, and extensively co-localizes with OCT4 and SOX2 throughout the mouse and human PS cell genome. While the core transcription factors define and govern pluripotency, in special circumstances PS cells can tolerate loss of SOX2 or NANOG or substitution with other factors, suggesting flexibility in pluripotency governance. Among the core

Table 1 | Different PS cell types and their developmental potentials

Starting cells	Pluripotent stem cell	Criteria for pluripotency						Ref.
		<i>In vitro</i> differentiation	Teratoma	Postnatal chimera	Germ line transmission	4n complementation	State of pluripotency	
Mouse germline tumour	ECCs	Yes	Yes	Yes	Yes	No	Naive	4
Mouse oocyte	Parthenogenetic ES cells	Yes	Yes	Yes	Yes	No	Naive	101
Mouse blastomere	ES cells	Yes	Yes	Yes	Yes	Yes	Naive	102
Mouse ICM	ES cells	Yes	Yes	Yes	Yes	Yes	Naive	5,6
Mouse Epiblast	EpiSCs	Yes	Yes	No	No		Primed	9,10
Mouse primordial germ cell	Embryonic germ cells	Yes	Yes	Yes	Yes	?	Naive	11,103
Mouse SSCs	GS cells, gPS cells; MASC	GS cells, gPS cells, MASC	GS cells, gPS cells, MASC	GS cells, gPS cells	GS cells, gPS cells	?	Naive (GS cells, gPS cells) Primed (MASCs)	12,104
Mouse somatic cells	iPS cells	Yes	Yes	Yes	Yes	Yes	Naive	3
Mouse somatic cells	Nuclear-transfer ES cells	Yes	Yes	Yes	Yes	Yes	Naive	105
Human germline tumour	ECCs	Yes	Yes	No	No	No	?	106
Human oocyte	Parthenogenetic ES cells	Yes	Yes	N/A	N/A	N/A	Primed	107
Human blastomere	ES cells	Yes	Yes	N/A	N/A	N/A	Primed	108
Human ICM	ES cells	Yes	Yes	N/A	N/A	N/A	Primed	7
Human somatic cells	iPS cells	Yes	Yes	N/A	N/A	N/A	Primed	109
Human somatic cells	Nuclear transfer ES cells	Yes	Yes	N/A	N/A	N/A	Primed	95

All cells listed are able to differentiate *in vitro*. Mouse oocyte-derived, blastocyst-derived ES cells, primordial germ-cell-derived embryonic germ cells, embryonic carcinoma cells (ECCs), SSC-derived cells, and iPS cells are able to generate chimaeras and contribute to the germ line. N/A, not applicable; ?, unknown. GS cells, germline stem cells; MASC, multipotent adult spermatogonial-derived stem cells; gPS cells, germline-derived pluripotent stem cells.

transcription factors, OCT4 has proven most indispensable and remains the preeminent pluripotency factor.

Mapping of OSN targets supports a model of regulatory control whereby OSN sustains self-renewal while restricting differentiation. OSN cooperatively bind their own promoters, forming an interconnected auto-regulatory loop^{17,18}. OSN activate a substantial fraction of protein-coding, miRNA, and non-coding RNA genes in ES cells, while also occupying genes encoding lineage-specific regulators^{27,28}. The promoters of many lineage regulators harbour both active (H3K4me3) and repressive (H3K27me3) histone marks, a bivalent state thought to facilitate activation of development genes upon exit from pluripotency²⁹. The capacity of OSN to activate genes necessary for maintaining ES cells, while repressing lineage-specifying regulators, chiefly accounts for the dual hallmark features of self-renewal and differentiation potential.

While OCT4 and SOX2 are expressed in all PS cells, PS cells can be classified into different states of pluripotency based on a complement of diagnostic molecular signatures that delineate proximity to the pre-implantation ICM or post-implantation epiblast, respectively (Fig. 1c). In mice, four key distinctions amongst the various pluripotent states have been described to date: (1) X chromosome status in female cells; (2) global levels of DNA methylation; (3) Oct4 enhancer utilization; and (4) expression levels of a select group of regulators designated as 'naive' transcription factors: Klf4, Klf2, Esrrb, Tfcp2l1, Tbx3 and Gbx2 (refs 10, 26, 30–33). These naive transcription factors, along with Nanog, are expressed at low levels or are absent in primed PS cells and can reset primed PS cells in conjunction with naive pluripotency culture conditions. The capacity of 'naive' transcription factors to reset primed PS cells suggests a regulatory intersection between naive transcriptional circuitry and epigenetic resetting of the DNA methylome and X chromosome.

A molecular 'ground state' in mouse ES cells can be enforced by cultivating cells in leukaemia inhibitory factor and small molecule inhibitors of Mek and Gsk3 kinases (2i/LIF conditions), which stabilizes the diagnostic signatures of pluripotency in the pre-implantation blastocyst^{30,34,35}. Ground state ES cells exhibit two active X chromosomes in female cells, low levels of DNA methylation, preferential utilization of the Oct4 distal enhancer, and naive transcription factor expression. In contrast, an alternative primed state is favoured by cultivation in FGF/ACTIVIN. Primed EpiSCs exhibit X-chromosome inactivation in female cells, high levels of DNA methylation, preferential utilization of the Oct4 proximal enhancer, and naive transcription factor repression. The molecular changes observed when ground state ES cells transition to primed

EpiSCs *in vitro* appear to mirror changes during maturation of pre-implantation epiblast to post-implantation epiblast *in vivo*^{14,36}.

Both naive and primed PS cells exhibit heterogeneity at the level of state markers and single cells, which we briefly discuss below. While serum-cultivated mouse PS cells form chimaeras capable of germline transmission (a functional hallmark of naive pluripotency), such PS cells also bear high DNA methylation levels reminiscent of post-implantation epiblast³⁰. EpiSCs also exhibit heterogeneity that can be altered via signalling pathway modulation. For example, region-specific EpiSCs (rsEpiSCs) preferentially engraft into posterior epiblasts and bear diagnostic markers of the post-implantation state, consistent with their status as primed PS cells³⁷. Yet, rsEpiSCs possess higher cloning efficiency, a feature typically associated with naive PS cells. Thus, like serum-cultivated ES cells, rsEpiSCs manifest features associated with different phases of pluripotency. Cumulatively, these observations suggest mouse pluripotency encompasses a spectrum of functional and molecular states, highlighting the imprecision of nomenclature in the face of biological complexity.

A caveat to the concept of ground state PS cells arises from single-cell studies suggesting inherent metastability in PS cells. Heterogeneous single-cell gene expression profiles, flow cytometry, and replating experiments indicate the coexistence of distinct molecular and functional states in serum-cultivated mouse ES cells³⁸. Even individual cells in more homogeneous ground state cultures have been reported to exhibit variable pluripotency transcription factor expression³⁹ and, while the origin and consequence of such heterogeneity are yet to be elucidated, the dynamic nature of pluripotency cannot be disregarded when classifying PS cell states. The markers distinguishing ground state from alternative PS cells remain relevant for evaluating novel PS cell types, especially claims of ground state human PS cells.

Adding additional nuance to the definitions of pluripotency, functional and molecular states are not always correlated. Mouse PS cells maintain molecular features of pluripotency, including expression of the core transcription factors, even when DNA methylation and H3K27 methylation are ablated^{40–43}, but cannot differentiate, and thereby lack functional pluripotency⁴⁴. Thus, while molecular signatures can suggest pluripotency, only functional tests can establish the true developmental potential of a cell. Unlike mouse ES cells, conventional 'primed' human ES cells cannot tolerate *DNMT1* deletion, emphasizing the functional differences between mouse and human ES cells, which we discuss in detail below⁴⁵. The observation that naive cells tolerate depletion of epigenetic regulators supports the concept of naive pluripotency as a

configuration with a reduced requirement for epigenetic repression compared to primed PS cells and somatic cells.

Functional assessment of pluripotency

A range of assays can be employed to reveal the developmental potential of PS cells: (1) *in vitro* differentiation; (2) teratoma formation; (3) chimaera formation; (4) germline transmission; (5) tetraploid complementation; and (6) single-cell chimaera formation. A summary of these assays along with their advantages and disadvantages is provided in Box 1.

In vitro differentiation to derivatives of all three embryonic germ layers—ectoderm, mesoderm and endoderm—represents the lowest hurdle for establishing pluripotency. Typically, culture conditions that maintain pluripotency are replaced by cocktails of differentiation-inducing cytokines, morphogens or chemicals, and markers of specific target tissues are then surveyed.

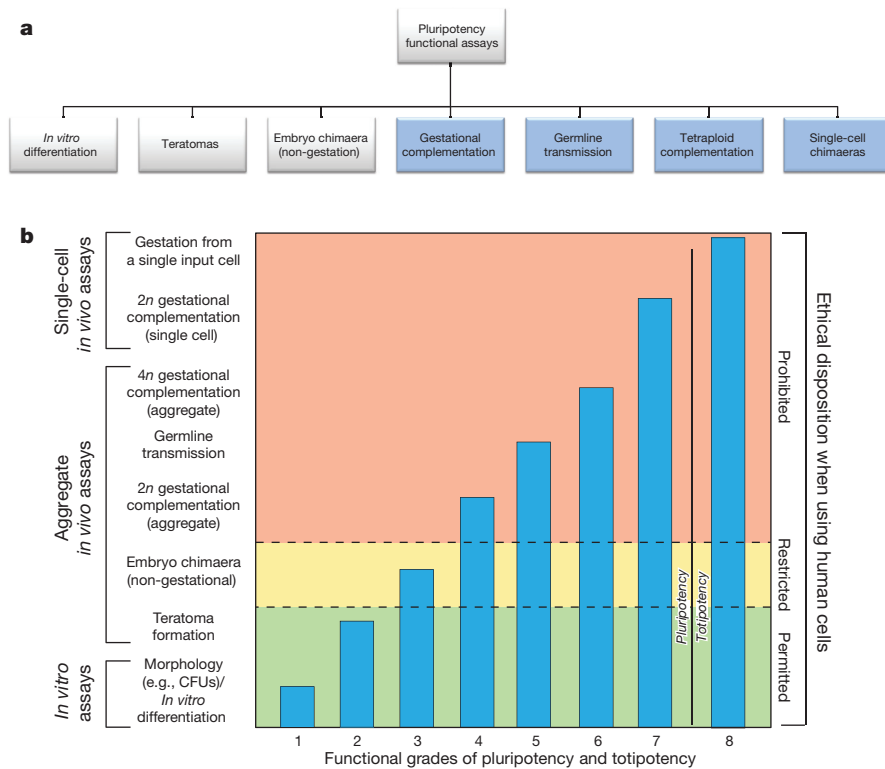
The teratoma formation assay assesses the spontaneous generation of differentiated tissues from the three germ layers following the injection of cells into immune-compromised mice. Histologic analysis of terato-

mas is neither quantitative nor capable of assessing every possible cell type. Incompletely reprogrammed cells can generate masses that superficially resemble teratomas yet lack terminal three-germ-layer differentiation, potentially leading to misinterpretation⁴⁶. Moreover, co-injection with matrices or scaffolds can elicit inflammatory or foreign-body reactions that can be misinterpreted as evidence of tissue differentiation, necessitating the use of lineage tracing or marker analysis to distinguish donor cells from reactive host tissue⁹³. Because teratomas are not generated from single cells, the teratoma assay assesses developmental potency at a population-based level.

A third differentiation assay, blastocyst chimaera formation, measures whether test cells can re-enter development when introduced into host embryos at either of two pre-implantation stages: by aggregation with cleavage-stage morulas or by injection into blastocysts⁴⁷. High-quality PS cells support normal development and generate high-grade chimaeras with extensive colonization of all embryonic tissues including the germ line, whereas less-potent PS cells produce either low chimaerism or reduced embryo viability.

BOX 1

Functional assays for pluripotency.



a, Overview of functional tests to assess developmental potency of PS cells. Blue boxes indicate assays that are restricted using human cells.

b, Functional assays for pluripotency, their grades of functional stringency, and ethical permissibility when using human cells. Analysis of *in vitro* characteristics, such as self-renewal capacity, colony morphology (CFU, colony-forming unit), and differentiation capacity *in vitro*, comprise a basic layer of pluripotency characterization. *In vivo* assays that measure differentiation capacity are taken as more robust indicators of potency.

Mouse PS-cell potency evaluation includes aggregate *in vivo* assays (that is, teratoma formation, embryo chimaeras (non-gestation), germline transmission, 2n/4n gestational complementation) and single-cell *in vivo* assays (that is, single-cell chimaeras and single-cell input gestations). 4n tetraploid complementation and single-cell chimaera formation are taken as more stringent functional assays for pluripotency.

The teratoma assay is the gold standard functional assay for assessing human PS-cell developmental potential. Chimaerism assays of human PS cells in murine embryos, as well as formation of primary human embryo chimaeras (non-gestation), are permissible under international stem-cell research guidelines¹¹⁰ after rigorous scientific and ethical review. Potency evaluation of primary human chimaeras by *in vivo* gestational complementation in humans is ethically impermissible.

The assays for totipotency are: (1) gestation from a single input cell; and (2) gestational complementation experiments from a single cell that demonstrate contribution to all tissues of the body and high-grade placenta contribution. Note that it is not necessarily the case that if a test cell performs well in a more stringent test, that it will definitely pass a less stringent test. For example, it is unclear if totipotent cells form teratomas.

A fourth assay, germline transmission, entails breeding chimaeras to produce all-donor PS cell-derived offspring, which thus demonstrates the capacity of test cells to generate functional gametes. The integration of donor cells into all tissues of viable late-stage embryos, postnatal or adult mice, followed by germline transmission, is a robust indicator of chromosomal integrity and of functional pluripotency.

A fifth assay applied to mouse cells, tetraploid complementation, measures the capacity of test PS cells to direct development of an entire organism. Donor PS cells are introduced into tetraploid ($4n$) host blastocysts, which are generated by electrofusion of blastomeres at the two-cell stage. Because $4n$ blastocysts cannot sustain normal embryonic development beyond mid-gestation, while tetraploid extra-embryonic tissues develop normally and support donor cells⁴⁸, any resulting embryos are derived essentially entirely from donor PS cells.

A sixth, highly stringent assay is to inject single-donor mouse PS cells into a morula or blastocyst⁴⁹. Genuine pluripotency is a property of a single cell and therefore chimaeras with widespread contribution from a single injected cell provide the clarity of clonal analysis. Both single-cell chimaerism and tetraploid complementation assays suffer from higher failure rates, but can be interpreted as the most definitive ways of demonstrating pluripotency.

Finally, while primed EpiSCs generate tri-lineage differentiation *in vitro* and form teratomas, EpiSCs rarely form chimaeras upon introduction into pre-implantation blastocysts. However, EpiSCs contribute to all germ layers when introduced into early post-implantation embryos in whole-embryo culture^{37,50}, although pluripotency of single cells has not yet been demonstrated.

Human pluripotent stem cells

Conventional human PS cells exhibit molecular hallmarks of primed state pluripotency, including preferential utilization of the OCT4 proximal enhancer, pronounced levels of DNA methylation, and a propensity for X chromosome inactivation in female cell lines⁵¹. Reports of human naive PS cells prompted some groups to attempt to assess potency by blastocyst chimaerism^{52–54}, constrained by the widespread acceptance that culture of human embryos for more than 14 days of development *in vitro*, or past the point of primitive streak formation (whichever is first), is ethically impermissible. Nevertheless, both primed and altered human PS cells have been introduced into mouse pre-implantation embryos^{52–55}. Human naive PS cells engraft into the mouse ICM^{52,54}, although contribution to cross-species chimaeras has been minimal⁵² or not detectable^{53,54}. By contrast, region-specific human PS cells engraft into the posterior epiblast of cultured murine post-implantation embryos, indicating limited cross-species chimaerism³⁷.

More compelling evidence for cross-species blastocyst chimaerism has been reported following injection of primate naive iPS cells into mouse blastocysts, leading to clonal contribution to solid tissues⁵⁶. Whereas primate ICM cells have thus far failed to form blastocyst chimaeras, unlike mouse ICM cells¹⁶, aggregation of primate blastomeres (totipotent cells) does produce chimaerism¹⁶. Nonetheless, a recent study described altered primate PS cells that can incorporate into host embryos and develop into chimaeric fetuses with low-grade contribution to all three germ layers and early germ cell progenitors⁵⁷. As in mice, high-grade contribution and germline transmission remain as more stringent tests to demonstrate naive pluripotency in primate ES cells.

Given the distinct behaviour of primate PS cells in chimaera studies, and lingering uncertainties about interspecies chimaerism, injecting human cells into mouse embryos needs additional validation before being accepted as a routine assay for stem-cell potency. Lacking robust functional assays for human stem-cell potency, transcriptional and epigenetic similarity of hypothetical ground state PS cells to the pluripotent cells in human pre-implantation embryos will remain the molecular standard for designation of human ground state PS cells (Fig. 1).

Erasure and resetting of DNA methylation is a molecular hallmark in mammalian pre-implantation and germline development. Human pre-implantation embryos have hypomethylated genomes. In contrast, ICM

outgrowths undergo genomic remethylation and established human ES cells maintain pronounced DNA hypermethylation, similar to mouse primed PS cells^{58,59}. Such epigenetic resetting appears to be controlled by a unique regulatory network present in pre-implantation embryos and the germ line. KLF4, TFCEP2L1, ESRRB, TBX3 and GBX2, transcription factors implicated in mouse naive pluripotency, have been detected in human pre-implantation epiblast and are transcriptionally repressed in derived human ES cells, similarly to mouse EpiSCs⁶⁰. However, the transcripts of certain murine naive transcription factors, such as KLF2, have not been detected in the human pre-implantation epiblast, revealing complexity. Additional species-specific differences also remain unresolved. The timing of X chromosome inactivation in human embryos is contentious^{61,62} and ‘epigenetic erosion’ of the X chromosome in primed human ES cells complicates our understanding of X chromosome regulation^{63,64}. Therefore, by current standards, we identify human ground state or naive PS cells according to molecular criteria used to delineate mouse ground state pluripotency, accepting that these criteria are tentative and subject to revision.

Acknowledging such caveats, a growing number of studies have demonstrated the feasibility of altering human PS cells towards a ‘metastable’ naive state of pluripotency^{52,65–67}. More convincingly, PS cells generated by the Jaenisch and Smith laboratories express transcription factors implicated in the governance of mouse ground state ES cells^{53,54}. While the X chromosome was inactive in human PS cells generated in the Jaenisch laboratory, we note again the uncertain significance of X chromosome status in human pluripotency^{52,53,61–64}. Cells ‘reset’ in the Smith laboratory exhibit a meaningful reduction in DNA methylation to levels approaching human pre-implantation embryos. However, the unclear activation of the OCT4 distal enhancer, and lack of detailed characterization of transgene-independent cell lines leaves open the question of whether the reset state is stable⁵⁴.

More experimental understanding of the transition from totipotency to pluripotency in the intact human or primate embryo will be needed to truly define the human ground state PS cell. Direct derivation of ground state ES cells from human embryos would be a landmark, highlighting the continued relevance of human ES cell research.

Potency in native somatic cells

As an organism progresses from the earliest embryonic stages to adulthood its cells become progressively restricted in developmental potency, and acquire epigenetic modifications that present barriers to dedifferentiation. However, germ cells, responsible for perpetuating the species, retain a unique chromatin state receptive to reprogramming to a naive pluripotent state by signalling pathway modulation alone. Cultivation of primordial germ cells in 2i/LIF, among other culture conditions, generates chimaera-competent naive pluripotent cells⁶⁸.

By contrast, acquisition of naive pluripotency from somatic cells requires the prolonged, combinatorial action of reprogramming transcription factors and ES cell growth conditions³. An exception to this principle is chemical reprogramming, suggesting that culture conditions alone can fully reverse the differentiated state to pluripotency⁶⁹. Notably, the final stage of chemical reprogramming is also induced by 2i/LIF. In contrast to mouse cells, our current capacity to generate human PS cells by signalling pathway modulation alone is more limited. The pluripotency of human embryonic germ cells and adult testis-derived human PS cells, both generated by culture of human germ cells, remains contentious, and small-molecule-based reprogramming of human somatic cells to pluripotency has not yet been demonstrated^{70–72}.

Alterations in cellular identity can accompany human disease. Chronic exposure to stomach acid from gastro-oesophageal reflux converts stratified squamous epithelium of the distal oesophagus to goblet-cell containing columnar epithelia more typical of the intestine, a condition termed Barrett’s oesophagus, which predisposes to adenocarcinoma. Metaplasia and other forms of tissue ectopias, where aberrant tissues form in unusual locations, suggest cell identity conversion occurs in the body. Thus it is intriguing to consider various claims of pluripotency for cells isolated from

perinatal or somatic tissues, such as multipotent adult progenitor cells^{73,74}, very small embryonic-like cells⁷⁵, multi-lineage differentiating stress-enduring cells⁷⁶, and endogenous pluripotent stem cells⁷⁷. When considering novel claims of expanded potency a strict criterion is demonstration that a single cell can differentiate into different cell types, a standard of clonal analysis lacking in most studies.

Evaluating totipotency features

A robust, bidirectional capacity to form both embryonic lineages and extra-embryonic trophoblast layers of the placenta, as well as yolk sac derivatives, distinguishes totipotency from pluripotency. While somatic cells are reset to totipotency following nuclear transfer into oocytes, to date no lab (to our knowledge) has claimed to propagate *in vitro* cells with totipotency equivalent to zygotes or blastomeres. Below, we briefly review previous claims of placental differentiation capacity in PS cells and propose how one might evaluate claims of totipotency (Table 2).

The most stringent demonstration of totipotency requires that a single cell produce a term birth under experimental conditions^{78–81}, a standard achieved in rodents and in non-human primates for single blastomeres extracted from pre-implantation embryos^{1,81}. Later-stage blastomeres may contribute to all embryonic and extra-embryonic tissues, and yet fail to support a viable conceptus because of reduced cell numbers at the blastocyst stage. Thus, an alternate and less stringent test of totipotency is the potential of genetically marked single cells to contribute extensively to both embryonic and extra-embryonic lineages after introducing donor cells into pre-blastocyst-stage embryos. In the mouse, for example, only isolated two-cell blastomeres can generate an entire conceptus⁸², but single blastomeres at the eight-cell stage still manifest totipotency in aggregation chimaeras¹. Sister blastomeres of a four-cell stage human embryo can develop individually into blastocysts with ICM and trophoblast cells⁸³. An essential feature of these functional tests of totipotency is demonstration of developmental capacity at the single-cell level.

Mouse PS cells with bidirectional developmental capacity for extra-embryonic and somatic fates have been claimed following specific genetic (for example, *Dnmt1* knockout⁸⁴) or cell culture modifications (for example, ground state^{39,85}) (summarized in Table 2). '*In vivo* reprogrammed' iPS cells purportedly contribute to the placenta, unlike ES cells or *in vitro* reprogrammed iPS cells⁸⁶. These studies reported differentiation into trophoblast-stem-like cells and the formation of blastocyst-like structures. However, the *in vivo* chimaera potential of trophoblast-stem-like cells was not assessed. Further, single cells did not yield robust high-grade contribution to the placenta³⁹. Thus, the definitive functional criterion for establishing totipotency, single-cell contribution to the trophoblast and ICM lineages, has not yet been demonstrated. The molecular changes associated with acquisition of totipotent-like developmental potential have differed across studies and include the expression of '2C-specific' genes, morula-specific genes, and extra-embryonic transcription factors. Therefore, by current

standards, accepting that the relevance of these molecular criteria are tentative and subject to revision, the essential criterion of totipotency remains functional, whereby a single cell generates both ICM and trophoblast fates in a transplantation assay. Ideally, detailed assessment of embryonic and extra-embryonic tissues should be made late in gestation, so that extensive and functional contribution can be demonstrated.

Conventional primed human PS cells reportedly form both trophoblast and primitive endoderm-like derivatives *in vitro*⁸⁷. However, confirmation of the identity of these derivatives has proven challenging⁸⁸. Injection of human naive PS cells into mouse embryos has not resulted in contribution to ICM and trophoblast lineages. Future claims of mouse totipotent stem cells will require stringent functional and molecular validation, while in humans, molecular criteria and comparison to primate species will have to suffice to establish plausibility.

Assessing provenance and potency via genomics

Advanced sequencing platforms have allowed researchers to generate a multitude of genomic and epigenomic data (for example, RNA sequencing (RNA-seq), chromatin immunoprecipitation sequencing (ChIP-seq) and bisulfite sequencing), enabling a more comprehensive description of cellular identity. Systems-level analyses have confirmed that direct reprogramming of somatic cells largely re-establishes molecular signatures associated with ES cells^{89,90}. These analyses also detected low-fidelity reprogramming, such as in intermediates and cells with epigenetic memory^{89,90}. Recently, genomic analyses have proven instrumental in defining ground state pluripotency. Thus, while not required for routine characterization of PS cells, genomic analyses play a critical role for benchmarking novel claims of reprogramming and PS cells (Box 2).

DNA sequencing also provides genetic fingerprints that can eliminate cell contamination as a confounder of reported results. Because cell line contamination is widespread, applying such genotyping methods to confirm cell line provenance is appropriate⁹¹. In the case of the STAP cell phenomenon, the authors reported acid-reprogrammed PS cells with features of totipotency. Our re-analysis of genomic data revealed unexpected mismatches in sex and genotype between donor somatic cells and converted STAP cells⁹². Further analysis of a STAP-derived cell line, Fgf4-induced stem cells, revealed a mixture that contained trophoblast stem cells, explaining the high-grade placenta colonization reported for Fgf4-induced stem cells. These findings are consistent with and extend the results of an extensive whole-genome sequencing analysis of STAP-related samples for the RIKEN investigation⁹³, which found contamination of purported STAP stem-cell lines with embryonic stem cells of a different genetic background⁹⁴.

By contrast, forensic genomics applied to sequencing data from two reports of nuclear-transfer-derived human ES cells (NT-hESCs) have confirmed cell line provenance^{95,96}. Inferred genome-wide single nuc-

Table 2 | Stem cells with reported bidirectional developmental potential

Cells	Criteria for totipotency							Ref.
	Genetic manipulation	Embryonic contribution	Placenta contribution	Yolk sac contribution	<i>In vitro</i> differentiation into trophoblast	Trophoblast stem-cell derivation	Single-cell injection	
<i>Dnmt1</i> KO ES cells	<i>Dnmt1</i> KO ES cells	?	Yes	?	Yes	No	No	84
'2C' ES cells/Kdm1a ES cells	2C reporter	Yes	Yes	Yes	Not tested/No	Not tested/No	No	85
Hex + 2i ES cells	Hex reporter	Yes	Yes	Yes	Cdx2 ⁺ trophoblast differentiation	Not tested/No	Yes	39
<i>In vivo</i> reprogrammed iPS cells	OSKM cassette	Yes	Yes	Yes	Yes	TSC-like cells, but no <i>in vivo</i> chimaeric placenta	No	86

KO, knockout; TSC, trophoblast stem cell; ?, unknown.

BOX 2

Forensic genomics for potency and provenance of PS cells.

Evaluating potency via genomic analyses.

Transcriptome analysis. Computational analysis can quantify the extent to which an experimental protocol converts a parental cell towards the target cell. For example, PluriTest¹¹¹ defines a pluripotency-specific signature based on a compendium of expression data sets from pluripotent and non-pluripotent cells and evaluates the presence of this signature in a given sample. CellNet is a bioinformatics algorithm that assesses the fidelity of cell fate conversion using cell- and tissue-specific gene regulatory networks¹¹².

Epigenomic characterization. Genome-wide profiling of chromatin features (for example, histone modifications, transcription factor binding, DNase I hypersensitivity, and DNA methylation) captures the epigenetic landscape of PS cells and the transitions that occur during reprogramming or differentiation. Combining gene expression and epigenetic maps can provide mechanistic insights into the fidelity of reprogrammed PS cells. For example, some mouse iPS cells fail to re-establish bivalent domains at developmental loci, which might reduce developmental competence for all tissue types⁸⁹. This failure could only be detected via epigenetic analysis.

Genomic integrity. Whole-genome sequencing (WGS) allows comprehensive identification of single nucleotide polymorphisms (SNPs) and a wide range of structural variations including copy number variants, copy-neutral events (such as translocations and inversions), and viral insertions, at base-pair resolution. Comparison of genomic variants before and after reprogramming will locate genomic alterations that may occur during reprogramming and potentially impact cellular function.

Genotyping for cell line provenance and contamination.

Provenance. Comparing the genotype of reprogrammed PS cells to parental cells enables verification of provenance. Genome-wide SNP arrays characterize a known set of SNPs and are sufficient for matching two samples. Based on the intensities of the probe hybridization reaction for each SNP and the ratio of the intensities between the two alleles, it is possible to estimate allele-specific copy numbers in addition to SNP genotypes. Sequencing-based assays such as exome or whole-genome sequencing can provide a more comprehensive characterization of SNPs; genotype information can also be inferred from RNA-seq and possibly other functional genomics data. Analyses of other types of genome variation such as microsatellites can also be used as a form of genetic fingerprinting. Microsatellite profiling, for example, is recommended by the International Cell Line Authentication Committee (ICLAC) for cultured cell lines¹¹³.

Contamination. Genome-wide SNP data can also be used to examine genetic heterogeneity of cell line cultures and to detect contamination with another cell line. For a homogeneous population, we expect to see sharp allele frequency (alternative over reference allele frequency ratio) distributions with a dominant peak near 0 (homozygous reference) and smaller peaks at ~0.5 (heterozygous) and 1 (homozygous alternative). When there is contamination by cells from different individuals or strains, we expect to observe small peaks at low allele frequencies (for example, at 0.05 and 0.1 if there is 10% contamination), corresponding to the alternative alleles from the second population of cells. With sequencing data, these estimates can be derived with greater precision, using both annotated SNPs and novel single-nucleotide variants. Sample contamination was detected in the STAP data with this analysis.⁹³

leotide variants (SNVs) from exome sequencing data classified samples generated in the Egli laboratory as genetically similar or dissimilar (Fig. 2). Parental donor fibroblasts and NT-hESCs possessed similar

SNV profiles, consistent with nuclear transfer origin. Independently sourced *in vitro*-fertilization-derived ES cells and parthenogenetic ES cells manifest distinct genetic provenance from parental donor fibroblasts and NT-hESCs, as expected. SNV genotyping also confirmed previously reported patterns of recombination in human parthenogenetic ES cells, concordant with observations in mouse parthenogenetic ES cells^{97,98}. Matching genotypes between parental fibroblasts and reprogrammed NT-hESCs were also confirmed in RNA-seq data generated in the Mitalipov laboratory (Supplementary Fig. 1). Collectively, these analyses support appropriate provenance of NT-hESCs and exclude a parthenogenetic origin for NT-hESCs.

Reproducibility of computational analyses

As genomic analyses can validate the provenance and confirm molecular signatures of novel PS cells, we advocate posting of relevant genomic data, metadata, and full details of computational analysis upon manuscript publication. Deposition of sequencing data to public repositories such as the Gene Expression Omnibus (GEO; <http://www.ncbi.nlm.nih.gov/geo/>) and Short Read Archive (SRA; <http://www.ncbi.nlm.nih.gov/sra>) is required by most peer-reviewed journals, but enforcement of sharing policies is highly variable, and complicated by the complexities of experimental design and data. Consequently, verification that full data and associated metadata have been deposited often requires expertise and time beyond what is available during peer review. Greater compliance by the stem-cell community in depositing all relevant genomic data and metadata as well as consistent enforcement by journals will promote reproducibility of results. We also recommend deposition of 'intermediate' data, the key steps and results obtained in the data analysis process. For full reproducibility of computational analysis, we also advocate release of the computer code, through a supplementary website or open source code management tools. We note that genomic analysis and availability of data, metadata, and methods are especially important for novel claims of reprogramming and altered stem-cell states.

Conclusion and future prospects

Here, we articulate a consensus definition of pluripotency predicated on both functional assessments of differentiation potential and diagnostic molecular signatures. Such an integration of functional and molecular hallmarks of pluripotency provides for a robust set of criteria against which to validate claims of pluripotency achieved by novel experimental strategies. Given the central role of core transcription factors in reprogramming somatic cells and maintaining the pluripotent state, failure to observe ES-cell-like levels of these transcription factors in studies asserting functional pluripotency from novel sources should merit scepticism and should be accompanied by strong evidence for alternative gene regulatory networks and mechanisms that maintain the unique pluripotent state of the mammalian genome. Another example of uncoupling between molecular and functional hallmarks is a report that overexpression of cell adhesion molecules such as E-cadherin can endow primed PS cells with the capacity to chimaerize the pre-blastocyst, with no evidence of resetting to naive pluripotency⁹⁹. Conversely, recent reports suggesting that reprogramming transitions through a transient state that molecularly resembles naive pluripotency, but without functional hallmarks of naive pluripotency, might not comprise bona fide naive pluripotency¹⁰⁰. While most labs deriving PS cells for routine use need not employ the comprehensive set of assays reviewed here, claims of novel states of potency or new means of deriving PS cells necessitate more comprehensive characterization and documentation.

Documentation of PS-cell states that span the continuum between ground state pluripotency and primed pluripotency provokes the question of how to define the human ground state. Further, reports that human PS cells can be 'reset' imply the feasibility of generating PS cells with bona fide totipotency. Ultimately, refined molecular benchmarking of reprogramming and more predictable experimental capture of altered

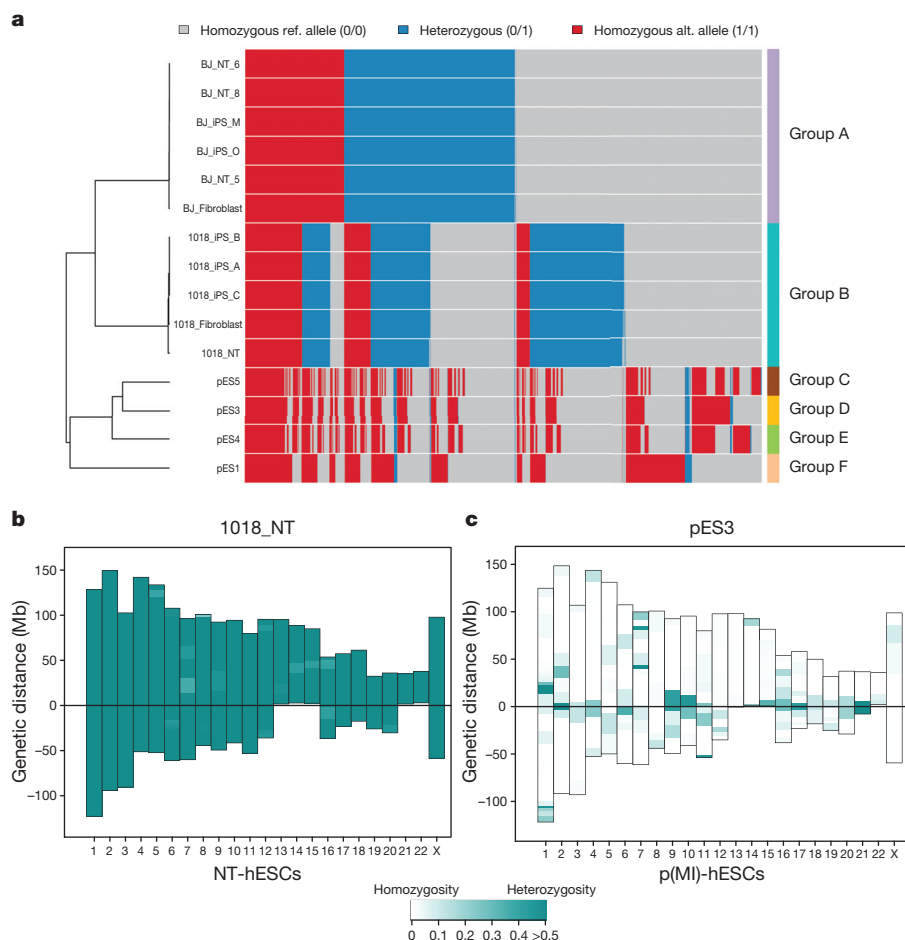


Figure 2 | Genomic provenance of nuclear transfer human embryonic stem cells (NT-hESCs). **a**, Single nucleotide sequence variants (SNVs) inferred using exome sequencing data using the human reference genome GRCh37. The selected SNVs are classified as homozygous for reference allele (0/0 genotype), homozygous for alternative allele (1/1 genotype) or heterozygous (0/1 genotype). Samples are clustered based on the sum of the edit distance between each SNV. The six different genotypes in three groups can be discerned: group A (BJ fibroblast and BJ fibroblast-reprogrammed human pluripotent stem-cell lines); group B (1018 fibroblast and 1018 fibroblast-reprogrammed human pluripotent stem-cell lines); and groups C–F (human parthenogenetic embryonic stem cells). **b**, Genome-wide SNP genotyping of a representative clone of NT-hESCs (Egli laboratory exome sequencing data) excluding parthenogenetic origin. Panels show genotypes for each chromosome, from centromere to telomere revealing blocks or haplotypes of markers. Mb, megabases. **c**, Genome-wide SNP genotyping of a representative clone of parthenogenetic (meiosis I) human embryonic stem cells (p(MI)-hES cells) (Egli laboratory exome sequencing data). Panels show genotypes for each chromosome, from centromere to telomere, revealing blocks or haplotypes of markers. Pericentromeric heterozygosity is consistent with a meiosis I parthenogenetic ES cell.

pluripotent states requires a more sophisticated understanding of human pre-implantation development.

For lasting scientific impact, claims of reprogramming and altered states of pluripotency should be broadly applicable to more than one experimental model and be independently replicated by multiple laboratories. Before publication, we encourage that researchers claiming landmark reprogramming advances first demonstrate replication by independent laboratories and incorporate forensic genomic analyses to confirm appropriate cell provenance. Science is ultimately a self-correcting process where the scientific community plays a crucial and collective role.

Received 6 May; accepted 26 August 2015.

- Kelly, S. J. Studies of the developmental potential of 4- and 8-cell stage blastomeres. *J. Exp. Zool.* **200**, 365–376 (1977).
- Gurdon, J. B. The developmental capacity of nuclei taken from intestinal epithelium cells of feeding tadpoles. *J. Embryol. Exp. Morphol.* **10**, 622–640 (1962).
A pioneering study that demonstrated that somatic cells can be reset to an early embryonic state via nuclear transplantation into eggs.
- Takahashi, K. & Yamanaka, S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* **126**, 663–676 (2006).
The landmark paper establishing that four transcription factors can reprogram somatic cells to a pluripotent state.
- Stevens, L. C. Studies on transplantable testicular teratomas of strain 129 mice. *J. Natl. Cancer Inst.* **20**, 1257–1275 (1958).
- Evans, M. J. & Kaufman, M. H. Establishment in culture of pluripotent cells from mouse embryos. *Nature* **292**, 154–156 (1981).
A paper establishing that pluripotent stem cells can be isolated from mouse blastocysts and be propagated *in vitro* as continuously growing cell lines.
- Martin, G. R. Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. *Proc. Natl. Acad. Sci. USA* **78**, 7634–7638 (1981).
A paper establishing that pluripotent stem cells can be isolated from mouse blastocysts and be propagated *in vitro* as continuously growing cell lines.

- Thomson, J. A. *et al.* Embryonic stem cells derived from human blastocysts. *Science* **282**, 1145–1147 (1998).
The landmark paper establishing that pluripotent stem cells can be isolated from human blastocysts.
- Thomson, J. A. *et al.* Isolation of a primate embryonic stem cell line. *Proc. Natl. Acad. Sci. USA* **92**, 7844–7848 (1995).
- Brons, I. G. *et al.* Derivation of pluripotent epiblast stem cells from mammalian embryos. *Nature* **448**, 191–195 (2007).
One of two pioneering studies that established that an alternative pluripotent state can be isolated from post-implantation mouse embryos resembling conventional human ES cells, suggesting that human ES cells might correspond to a post-implantation state.
- Tesar, P. J. *et al.* New cell lines from mouse epiblast share defining features with human embryonic stem cells. *Nature* **448**, 196–199 (2007).
One of two pioneering studies that established that an alternative pluripotent state from post-implantation mouse embryos that resembles conventional human ES cells, suggesting that human ES cells might correspond to a post-implantation state.
- Matsui, Y., Zsebo, K. & Hogan, B. L. M. Derivation of pluripotent embryonic stem cells from murine primordial germ cells in culture. *Cell* **70**, 841–847 (1992).
- Kanatsu-Shinohara, M. *et al.* Generation of pluripotent stem cells from neonatal mouse testis. *Cell* **119**, 1001–1012 (2004).
- Ko, K. *et al.* Induction of pluripotency in adult unipotent germline stem cells. *Cell Stem Cell* **5**, 87–96 (2009).
- Boroviak, T. *et al.* The ability of inner-cell-mass cells to self-renew as embryonic stem cells following epiblast specification. *Nature Cell Bio.* **16**, 516–528 (2014).
- Nichols, J. & Smith, A. Naive and primed pluripotent states. *Cell Stem Cell* **4**, 487–492 (2009).
- Tachibana, M. *et al.* Generation of chimeric rhesus monkeys. *Cell* **148**, 285–295 (2012).
- Boyer, L. A. *et al.* Core regulatory circuitry in human embryonic stem cells. *Cell* **122**, 947–956 (2005).
A study that affirmed the principle that OCT4, SOX2, and NANOG constitute a core regulatory circuitry that explains the self-renewal and differentiation capacity of PS cells.
- Loh, Y. H. *et al.* The Oct4 and Nanog transcription network regulates pluripotency in mouse embryonic stem cells. *Nature Genet.* **38**, 431–440 (2006).
- Nichols, J. *et al.* Formation of pluripotent stem cells in the mammalian embryo depends on the POU transcription factor Oct4. *Cell* **95**, 379–391 (1998).

20. Chambers, I. *et al.* Functional expression cloning of Nanog, a pluripotency sustaining factor in embryonic stem cells. *Cell* **113**, 643–655 (2003).
21. Mitsui, K. *et al.* The homeoprotein Nanog is required for maintenance of pluripotency in mouse epiblast and ES cells. *Cell* **113**, 631–642 (2003).
22. Avilion, A. *et al.* Multipotent cell lineages in early mouse development depend on SOX2 function. *Genes Dev.* **17**, 126–140 (2003).
23. Masui, S. *et al.* Pluripotency governed by Sox2 via regulation of Oct3/4 expression in mouse embryonic stem cells. *Nature Cell Biol.* **9**, 625–635 (2007).
24. Wang, Z., Oron, E., Nelson, B., Razis, S. & Ivanova, N. Distinct lineage specification roles for NANOG, OCT4, and SOX2 in human embryonic stem cells. *Cell Stem Cell* **10**, 440–454 (2012).
25. Chambers, I. *et al.* Nanog safeguards pluripotency and mediates germline development. *Nature* **450**, 1230–1234 (2007).
26. Silva, J. *et al.* Nanog is the gateway to the pluripotent ground state. *Cell* **138**, 722–737 (2009).
27. Chen, X. *et al.* Integration of external signaling pathway with the core transcriptional network in embryonic stem cells. *Cell* **133**, 1106–1117 (2008).
28. Kim, J., Chu, J., Shen, X., Wang, J. & Orkin, S. H. An extended transcriptional network for pluripotency of embryonic stem cells. *Cell* **132**, 1049–1061 (2008).
29. Bernstein, B. E. *et al.* A bivalent chromatin structure marks key developmental genes in embryonic stem cells. *Cell* **125**, 315–326 (2006).
30. Leitch, H. G. *et al.* Naive pluripotency is associated with global DNA hypomethylation. *Nature Struct. Mol. Biol.* **20**, 311–316 (2013).
31. Guo, G. *et al.* Klf4 reverts developmentally programmed restriction of ground state pluripotency. *Development* **136**, 1063–1069 (2009).
- The first study to describe the association between 2i cultivation and DNA hypomethylation, linking in vitro naive pluripotency with the DNA hypomethylation observed in pre-implantation embryos and the germ line.**
31. Guo, G. *et al.* Klf4 reverts developmentally programmed restriction of ground state pluripotency. *Development* **136**, 1063–1069 (2009).
- The first paper to convert EpiSCs to naive pluripotency, affirming the concept that the naive and primed states of pluripotency are interconvertible.**
32. Bao, S. *et al.* Epigenetic reversion of post-implantation epiblast to pluripotent embryonic stem cells. *Nature* **461**, 1292–1295 (2009).
33. Dunn, S. J., Martello, G., Yordanov, B., Emmott, S. & Smith, A. G. Defining an essential transcription factor program for naive pluripotency. *Science* **344**, 1156–1160 (2014).
34. Ying, Q. L. *et al.* The ground state of embryonic stem cell self-renewal. *Nature* **453**, 519–523 (2008).
- The first study to report the remarkable synergism between MEK and GSK3 inhibition and the first articulation of the concept of a pluripotent 'ground state'.**
35. Marks, H. *et al.* The transcriptional and epigenomic foundations of ground state pluripotency. *Cell* **149**, 590–604 (2012).
36. Smith, Z. D. *et al.* A unique regulatory phase of DNA methylation in the early mammalian embryo. *Nature* **484**, 339–344 (2012).
37. Wu, J. *et al.* An alternative pluripotent state confers interspecies chimeric competency. *Nature* **521**, 316–321 (2015).
38. Kumar, R. M. *et al.* Deconstructing transcriptional heterogeneity in pluripotent stem cells. *Nature* **516**, 56–61 (2014).
39. Morgani, S. M. *et al.* Totipotent embryonic stem cells arise in ground-state culture conditions. *Cell Rep.* **3**, 1945–1957 (2013).
40. Chamberlain, S. J., Yee, D. & Magnuson, T. Polycomb repressive complex 2 is dispensable for maintenance of embryonic stem cell pluripotency. *Stem Cells* **26**, 1496–1505 (2008).
41. Okano, M., Bell, D. W., Haber, D. A. & Li, E. DNA methyltransferases Dnmt3a and Dnmt3b are essential for *de novo* methylation and mammalian development. *Cell* **99**, 247–257 (1999).
42. Li, E., Bestor, T. H. & Jaenisch, R. Target mutation of the DNA methyltransferase gene results in embryonic lethality. *Cell* **69**, 915–926 (1992).
43. Tsumura, A. *et al.* Maintenance of self-renewal ability of mouse embryonic stem cells in the absence of DNA methyltransferases Dnmt1, Dnmt3a, and Dnmt3b. *Genes Cells* **11**, 805–814 (2006).
44. Beard, C., Li, E. & Jaenisch, R. Loss of methylation activates XIST in somatic but not in embryonic cells. *Genes Dev.* **9**, 2325–2334 (1995).
45. Liao, J. *et al.* Targeted disruption of DNMT1, DNMT3A, and DNMT3B in human embryonic stem cells. *Nature Genet.* **47**, 469–478 (2015).
46. Chan, E. M. *et al.* Live cell imaging distinguishes bona fide human iPS cells from partially reprogrammed cells. *Nature Biotechnol.* **27**, 1033–1037 (2009).
47. Nagy, A., Gocza, E. *et al.* Embryonic stem cells alone are able to support fetal development in the mouse. *Development* **110**, 815–821 (1990).
48. Nagy, A., Rossant, J., Nagy, R., Abramow-Newerly, W. & Roder, J. C. Derivation of completely cell culture-derived mice from early-passage embryonic stem cells. *Proc. Natl Acad. Sci. USA* **90**, 8424–8428 (1993).
49. Wang, Z. & Jaenisch, R. At most three ES cells contribute to the somatic lineages of chimeric mice produced by ES-tetraploid complementation. *Dev. Biol.* **275**, 192–201 (2004).
50. Huang, Y., Osorno, R., Tsakiridis, A. & Wilson, V. *In vivo* differentiation potential of epiblast stem cells revealed by chimeric embryo formation. *Cell Rep.* **2**, 1571–1578 (2012).
51. Hanna, J. *et al.* Human embryonic stem cells with biological and epigenetic characteristics similar to those of mouse ESCs. *Proc. Natl Acad. Sci. USA* **107**, 9222–9227 (2010).
52. Gafni, O. *et al.* Derivation of novel human ground state naive pluripotent stem cells. *Nature* **504**, 282–286 (2013).
53. Theunissen, T. W. *et al.* Systematic identification of culture conditions for induction and maintenance of naive human pluripotency. *Cell Stem Cell* **15**, 471–487 (2014).
54. Takashima, Y. *et al.* Resetting transcription factor control circuitry toward ground-state pluripotency in human. *Cell* **158**, 1254–1269 (2014).
55. James, D., Noggle, S. A., Swigut, T. & Brivanlou, A. H. Contribution of human embryonic stem cells to mouse blastocyst. *Dev. Biol.* **295**, 90–102 (2006).
56. Fang, R. *et al.* Generation of naive induced pluripotent stem cells from rhesus monkey fibroblasts. *Cell Stem Cell* **15**, 488–496 (2014).
57. Chen, Y. *et al.* Generation of cynomolgus monkey chimeric fetuses using embryonic stem cells. *Cell Stem Cell* **17**, 116–124 (2015).
58. Guo, H. *et al.* The DNA methylation landscape of human early embryos. *Nature* **511**, 606–610 (2014).
59. Smith, Z. D. *et al.* DNA methylation dynamics of the human preimplantation embryo. *Nature* **511**, 611–615 (2014).
60. Yan, L. *et al.* Single-cell RNA-seq profiling of human preimplantation embryos and embryonic stem cells. *Nature Struct. Mol. Biol.* **20**, 1131–1139 (2013).
61. Okamoto, I. *et al.* Eutherian mammals use diverse strategies to initiate X-chromosome inactivation during development. *Nature* **472**, 370–374 (2011).
62. O'Leary, T. *et al.* Tracking the progression of the human inner cell mass during embryonic stem cell derivation. *Nature Biotechnol.* **30**, 278–282 (2012).
63. Silva, S. S., Rowntree, R. K., Mekhoubad, S. & Lee, J. T. X-chromosome inactivation and epigenetic fluidity in human embryonic stem cells. *Proc. Natl Acad. Sci. USA* **105**, 4820–4825 (2008).
64. Anguera, M. C. *et al.* Molecular signatures of human induced pluripotent stem cells highlight sex differences and cancer genes. *Cell Stem Cell* **11**, 75–90 (2012).
65. Chan, Y. S. *et al.* Induction of a human pluripotent state with distinct regulatory circuitry that resembles preimplantation epiblast. *Cell Stem Cell* **13**, 663–675 (2013).
66. Ware, C. B. *et al.* Derivation of naive human embryonic stem cells. *Proc. Natl Acad. Sci. USA* **111**, 4484–4489 (2014).
67. Valamehr, B. *et al.* Platform for induction and maintenance of transgene-free hiPSCs resembling ground state pluripotent stem cells. *Stem Cell Rep.* **2**, 366–381 (2014).
68. Leitch, H. G. *et al.* Rebuilding pluripotency from primordial germ cells. *Stem Cell Rep.* **1**, 66–78 (2013).
69. Hou, P. *et al.* Pluripotent stem cells induced from mouse somatic cells by small-molecule compounds. *Science* **341**, 651–654 (2013).
70. Shambhoo, M. J. *et al.* Derivation of pluripotent stem cells from cultured human primordial germ cells. *Proc. Natl Acad. Sci. USA* **95**, 13726–13731 (1998).
71. Conrad, S. *et al.* Generation of pluripotent stem cells from adult human testis. *Nature* **456**, 344–349 (2008); retraction **512**, 338 (2014).
72. Ko, K. *et al.* Human adult germline stem cells in question. *Nature* **465**, E1 (2010).
73. Reyes, M. & Verfaillie, C. M. Characterization of multipotent adult progenitor cells, a subpopulation of mesenchymal stem cells. *Ann. NY Acad. Sci.* **938**, 231–235 (2001).
74. Jiang, Y. *et al.* Pluripotency of mesenchymal stem cells derived from adult marrow. *Nature* **418**, 41–49 (2002).
75. Kucia, M. *et al.* A population of very small embryonic-like (VSEL) CXCR4⁺SSEA-1⁺Oct4⁺ stem cells identified in adult bone marrow. *Leukemia* **20**, 857–869 (2006).
76. Kuroda, Y. *et al.* Isolation, culture, and evaluation of multilineage-differentiating stress-enduring (Muse) cells. *Nature Protocols* **8**, 1391–1415 (2013).
77. Roy, S. *et al.* Rare somatic cells from human breast tissue exhibit extensive lineage plasticity. *Proc. Natl Acad. Sci. USA* **110**, 4598–4603 (2013).
78. Nicholas, J. & Hall, B. Experiments on developing rats: II. The development of isolated blastomeres and fused eggs. *J. Exp. Zool.* **90**, 441–459 (1942).
79. Tarkowski, A. K. Experiments on the development of isolated blastomeres of mouse eggs. *Nature* **184**, 1286–1287 (1959).
80. Willadsen, S. M. & Polge, C. Attempts to produce monozygotic quadruplets in cattle by blastomere separation. *Vet. Rec.* **108**, 211–213 (1981).
81. Mitalipov, S. M. *et al.* Monozygotic twinning in rhesus monkeys by manipulation of *in vitro*-derived embryos. *Biol. Reprod.* **66**, 1449–1455 (2002).
82. Rossant, J. Postimplantation development of blastomeres isolated from 4- and 8-cell mouse eggs. *J. Embryol. Exp. Morphol.* **36**, 283–290 (1976).
83. Van de Velde, H., Cauffman, G., Tournaye, H., Devroey, P. & Liebaers, I. The four blastomeres of a 4-cell stage human embryo are able to develop individually into blastocysts with inner cell mass and trophoblast. *Hum. Reprod.* **23**, 1742–1747 (2008).
84. Ng, R. K. *et al.* Epigenetic restriction of embryonic cell lineage fate by methylation of Elf5. *Nature Cell Biol.* **10**, 1280–1290 (2008).
85. Macfarlan, T. S. *et al.* Embryonic stem cell potency fluctuates with endogenous retrovirus activity. *Nature* **487**, 57–63 (2012).
86. Abad, M. *et al.* Reprogramming *in vivo* produces teratomas and iPS cells with totipotency features. *Nature* **502**, 340–345 (2013).
87. Xu, R. H. *et al.* BMP4 initiates human embryonic stem cell differentiation to trophoblast. *Nature Biotechnol.* **20**, 1261–1264 (2002).
88. Bernardo, A. S. *et al.* BRACHYURY and CDX2 mediate BMP-induced differentiation of human and mouse pluripotent stem cells into embryonic and extraembryonic lineages. *Cell Stem Cell* **9**, 144–155 (2011).
89. Mikkelsen, T. S. *et al.* Dissecting direct reprogramming through integrative genomic analysis. *Nature* **454**, 49–55 (2008).
90. Kim, K. *et al.* Epigenetic memory in induced pluripotent stem cells. *Nature* **467**, 285–290 (2010).
91. Yu, M. *et al.* A resource for cell line authentication, annotation, and quality control. *Nature* **520**, 307–311 (2015).
92. De Los Angeles, A. *et al.* Failure to replicate the STAP cell phenomenon. *Nature* <http://dx.doi.org/10.1038/nature15513> (this issue).
93. RIKEN. Report on STAP Cell Research Paper Investigation. <http://www3.riken.jp/stap/e/c13document52.pdf> (2014).

94. Konno, D. *et al.* STAP cells are derived from ES cells. *Nature* <http://dx.doi.org/10.1038/nature15366> (this issue).
95. Tachibana, M. *et al.* Human embryonic stem cells derived by somatic cell nuclear transfer. *Cell* **153**, 1228–1238 (2013).
This study was the first to demonstrate the feasibility of somatic cell nuclear transfer to reset human cells to totipotency.
96. Ma, H. *et al.* Human oocytes reprogram adult somatic nuclei of a type I diabetic to diploid pluripotent stem cells. *Nature* **511**, 177–183 (2014).
97. Kim, K. *et al.* Histocompatible embryonic stem cells by parthenogenesis. *Science* **315**, 482–486 (2007).
98. Kim, K. *et al.* Recombination signatures distinguish embryonic stem cells derived by parthenogenesis and somatic cell nuclear transfer. *Cell Stem Cell* **1**, 346–352 (2007).
99. Ohtsuka, S., Nishikawa-Torikai, S. & Niwa, H. E-cadherin promotes incorporation of mouse epiblast stem cells into normal development. *PLoS ONE* **7**, e45220 (2012).
100. Cacchiarelli, D. *et al.* Integrative analyses of human reprogramming reveal dynamic nature of induced pluripotency. *Cell* **162**, 412–424 (2015).
101. Narasimha, M., Barton, S. C. & Surani, M. A. The role of the paternal genome in the development of the mouse germ line. *Curr. Biol.* **7**, 881–884 (1997).
102. Wakayama, S. *et al.* Efficient establishment of mouse embryonic stem cell lines from single blastomeres and polar bodies. *Stem Cells* **25**, 986–993 (2007).
103. Resnick, J. L., Bixler, L. S., Cheng, L. & Donovan, P. J. Long-term proliferation of mouse germ cells in culture. *Nature* **359**, 550–551 (1992).
104. Seandel, M. *et al.* Generation of functional multipotent adult stem cells from GPR125⁺ germline progenitors. *Nature* **449**, 346–350 (2007).
105. Munsie, M. J. *et al.* Isolation of pluripotent embryonic stem cells from reprogrammed adult mouse somatic cell nuclei. *Curr. Biol.* **10**, 989–992 (2000).
106. Andrews, P. W., Bronson, D. L., Benham, F., Strickland, S. & Knwles, B. B. A comparative study of eight cell lines derived from human testicular teratocarcinoma. *Int. J. Cancer* **26**, 269–280 (1980).
107. Revazova, E. S. *et al.* Patient-specific stem cell lines derived from human parthenogenetic blastocysts. *Cloning Stem Cells* **9**, 432–439 (2007).
108. Klimanskaya, I., Chung, Y., Becker, S., Lu, S. J. & Lanza, R. Human embryonic stem cell lines derived from single blastomeres. *Nature* **444**, 481–485 (2006).
109. Takahashi, K. *et al.* Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell* **131**, 861–872 (2007).
110. ISSCR. Guidelines for the conduct of human embryonic stem cell research <http://www.isscr.org/docs/default-source/hesc-guidelines/isscrhescguidelines2006.pdf> (2006).
111. Müller, F.-J. *et al.* A bioinformatics assay for pluripotency in human cells. *Nature Methods* **8**, 315–317 (2011).
112. Cahan, P. *et al.* CellNet: network biology applied to stem cell engineering. *Cell* **158**, 903–915 (2014).
113. International Cell Line Authentication Committee. Guide to human cell line authentication (2012) http://standards.atcc.org/kwspub/home/the_international_cell_line_authentication_committee-iclac/_Authentication_SOP.pdf.

Supplementary Information is available in the online version of the paper.

Acknowledgements We would like to thank B. Johannesson and D. Egli for providing sequencing data from nuclear-transfer-derived human embryonic stem cells. We would also like to thank P. J. Tesar, S. Byrne, A. Urbach, Y.-H. Loh, R. Zhao, K. Tsankov, J. Powers, T. Schlaeger, L. Daheron, N. Shyh-Chang, Y. S. Chan, and other members of the Daley laboratory for critical reading of this manuscript. G.Q.D. is an Investigator of the Howard Hughes Medical Institute.

Author Contributions A.D.L.A. and G.Q.D. conceived the study and wrote the manuscript. F.F., R.X., S.L. and P.J.P. performed bioinformatic analyses and wrote the forensic genomics section. Y.F., N.B., H.D., K.H., R.J., H.G.L., M.W.L., E.L., D.P., J.R. and M.W. assisted with writing.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to G.Q.D. (george.daley@childrens.harvard.edu).

Epicardial FSTL1 reconstitution regenerates the adult mammalian heart

Ke Wei^{1,2*}, Vahid Serpooshan^{3*}, Cecilia Hurtado^{1,2}, Marta Diez-Cuñado^{1,2,3}, Mingming Zhao³, Sonomi Maruyama⁴, Wenhong Zhu^{1,2}, Giovanni Fajardo³, Michela Nosedà⁵, Kazuto Nakamura⁴, Xueying Tian⁶, Qiaozhen Liu⁶, Andrew Wang³, Yuka Matsuura³, Paul Bushway^{1,2}, Wenqing Cai^{1,2}, Alex Savchenko^{1,2}, Morteza Mahmoudi^{3,7}, Michael D. Schneider⁵, Maurice J. B. van den Hoff⁸, Manish J. Butte³, Phillip C. Yang³, Kenneth Walsh⁴, Bin Zhou^{6,9}, Daniel Bernstein³, Mark Mercola^{1,2} & Pilar Ruiz-Lozano³

The elucidation of factors that activate the regeneration of the adult mammalian heart is of major scientific and therapeutic importance. Here we found that epicardial cells contain a potent cardiogenic activity identified as follistatin-like 1 (Fstl1). Epicardial Fstl1 declines following myocardial infarction and is replaced by myocardial expression. Myocardial Fstl1 does not promote regeneration, either basally or upon transgenic overexpression. Application of the human Fstl1 protein (FSTL1) via an epicardial patch stimulates cell cycle entry and division of pre-existing cardiomyocytes, improving cardiac function and survival in mouse and swine models of myocardial infarction. The data suggest that the loss of epicardial FSTL1 is a maladaptive response to injury, and that its restoration would be an effective way to reverse myocardial death and remodelling following myocardial infarction in humans.

The epicardium of the heart is an external epithelial layer that contributes to myocardial growth during development by providing progenitor cells^{1,2} as well as mitogens, including FGFs, IGF2, and PDGFs^{3–5}. Recent studies suggest that the epicardium might also preserve function of the adult myocardium following injury, possibly as a source of myogenic progenitors^{6,7}. To our knowledge no epicardial-secreted factors have been shown to support adult myocardial regeneration in mammals to date.

Epicardial signal activates cardiomyocyte division

We co-cultured an epicardial mesothelial cell (EMC) line with Myh6⁺ mouse embryonic stem cell-derived cardiomyocytes (referred to as mCMs^{ESC}; Extended Data Fig. 1, Supplementary Videos 1 and 2, and Methods). Co-cultures consistently increased the number of cardiomyocytes (α -actinin⁺ cells, Fig. 1a–c) and the expression of cardiomyocyte markers (Fig. 1d). Conditioned media from EMC cultures recapitulated this effect (Fig. 1e–h). The number of α -actinin⁺ cells exhibiting rhythmic Ca²⁺ transients also increased with the addition of EMC media (8.6-fold) (Fig. 1i), as quantified automatically by kinetic imaging cytometry. Similarly, conditioned media prepared from adult epicardial-derived cells⁸ increased proliferation and nearly doubled the incidence of aurora B kinase in the cleavage furrow connecting adjacent embryonic cardiomyocytes (Tnnt2⁺ cells; 0.19 to 0.33%, $P < 0.05$, Fig. 1j–m), indicating a secreted activity in the adult epicardium that promotes cytokinesis of embryonic cardiomyocytes.

Engineered epicardium improves function after injury

We next evaluated the effect of epicardial-secreted factors in the adult injured heart by delivering conditioned media in three-dimensional

collagen nano-fibrillar patches⁹. Patches were designed with an elastic modulus emulating the embryonic epicardium ($E \sim 12$ kPa)⁹, lower than the mature epicardium ($E > 30$ –40 kPa) and fibrotic cardiac tissue ($E > 100$ kPa), but higher than those for the most currently used scaffolding biomaterials ($E \leq 1$ kPa) (Fig. 1n, o). Patches seeded with EMC-media (33% of total volume) were sutured onto the heart immediately following surgical-induced myocardial infarction (MI, permanent ligation of the left anterior descending LAD coronary artery, Fig. 1p, q). Two weeks later, patch-treated hearts (both with or without EMC-media) showed improved morphometric parameters (Fig. 1r–t and Extended Data Table 1), consistent with collagen patch providing a mechanical support that inhibits remodelling⁹. Notably, only patch with EMC media treatment improved cardiac function (Fig. 1r, s and Extended Data Table 1).

Fstl1 is an epicardial cardiomyogenic factor

To identify bioactive proteins, we analysed EMC-conditioned media by mass spectrometry. Comparison of spectra to the IPI rat database identified 1,596 peptide reads corresponding to 311 unique proteins. Ten secreted proteins with the highest spectral counts were selected for testing in the mCMs^{ESC} assay. Of these, cardiogenic activity was noted only with follistatin-like 1 (also known as Fstl1, FRP or TSC36, accession number: NP_077345.1) (Fig. 2a). Unlike follistatin, Fstl1 does not block activin and its biochemical and biological functions are poorly characterized¹⁰. Fstl1 levels increase in the blood stream following acute MI and, for this reason, it has been considered a biomarker for acute coronary syndrome¹¹.

Treating mCMs^{ESC} for 8 days with bacterially synthesized recombinant human FSTL1 (10 ng ml⁻¹) increased the number of

¹Department of Bioengineering, University of California, San Diego, La Jolla, California 92037, USA. ²Sanford-Burnham-Prebys Medical Discovery Institute, 10901 N. Torrey Pines Road, La Jolla, California 92037, USA. ³Stanford Cardiovascular Institute and Department of Pediatrics, Stanford University School of Medicine, 300 Pasteur Drive, Stanford, California 94305, USA. ⁴Whitaker Cardiovascular Institute, Boston University School of Medicine, Boston, Massachusetts 02118, USA. ⁵Imperial College London, Faculty of Medicine, Imperial Centre for Translational and Experimental Medicine, Du Cane Road, London W12 0NN, UK. ⁶Key Laboratory of Nutrition and Metabolism, Institute for Nutritional Sciences, and Shanghai Institutes for Biological Sciences, Graduate School of the Chinese Academy of Sciences, Chinese Academy of Sciences, Shanghai 200031, China. ⁷Nanotechnology Research Center, Faculty of Pharmacy, Tehran University of Medical Sciences, 1417613151 Tehran, Iran. ⁸Academic Medical Center, Dept Anatomy, Embryology and Physiology, Meibergdreef 15, 1105AZ Amsterdam, The Netherlands. ⁹CAS Center for Excellence in Brain Science, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, Shanghai 200031, China.

*These authors contributed equally to this work.

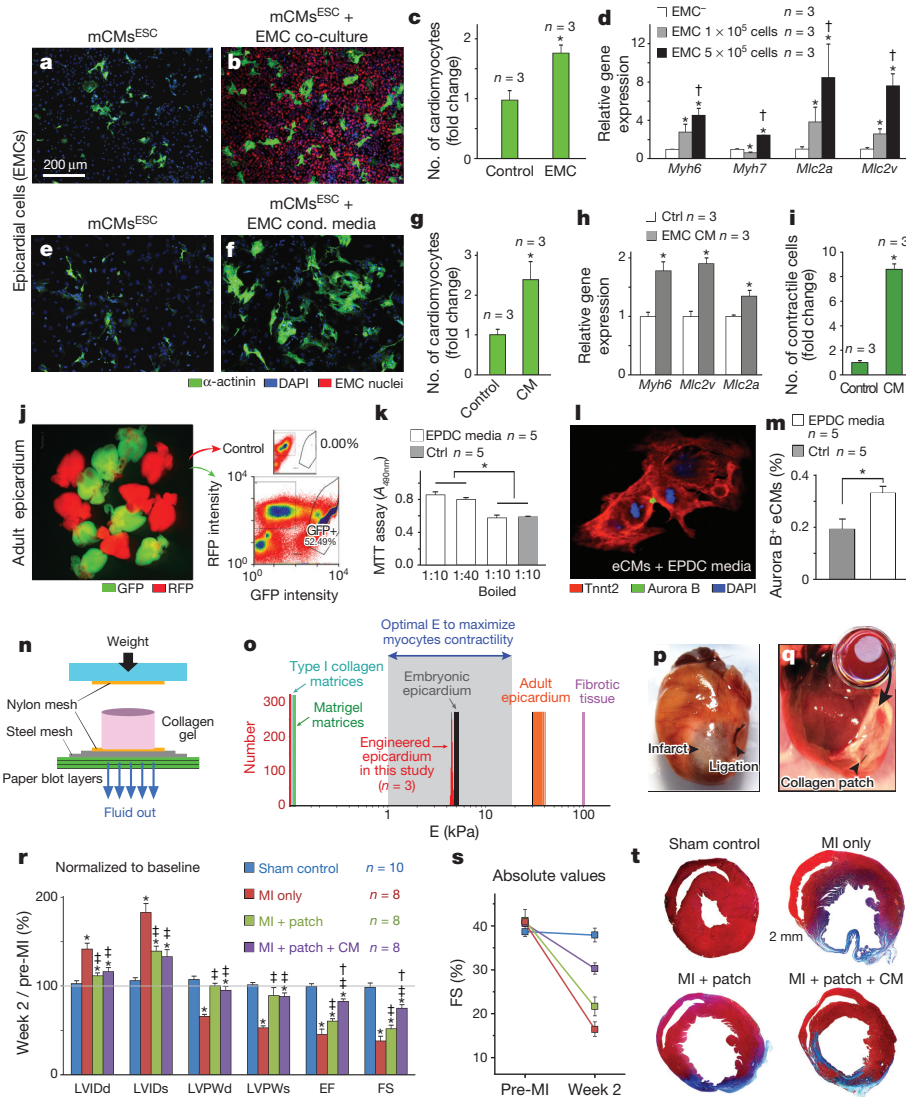


Figure 1 | Epicardial secretome has cardiogenic activity, and improves cardiac function after MI via embryonic epicardium-like patches. **a–d**, Co-culture of mCMEs^{ESC} cardiomyocytes with epicardial EMC cells. **a, b**, Representative micrographs. **c, d**, Quantification of myocyte number (c) and cardiac gene expression (d). * $P < 0.05$ compared to acellular (EMC[−]) control; † $P < 0.05$ compared to 1×10^5 cells condition. **e–i**, Culture of mCMEs^{ESC} cardiomyocytes with EMC-conditioned media. Representative micrographs (e, f). Quantification of myocyte number (g), cardiac gene expression (h), and cardiomyocytes with rhythmic calcium transients (i). * $P < 0.05$ compared to control. **j–m**, Effect of adult epicardial media on embryonic cardiomyocytes from E12.5 GFP⁺ cells (*Tnnt2-cre; Rosa26^{mTmG/+}*) (j). Conditioned media obtained from adult epicardial-derived cells (EPDCs) promotes cardiomyocyte proliferation that can be heat-inactivated (k) and

cytokinesis analysed by double immunostaining for aurora B and Tnnt2 (cardiomyocytes) (l, m). * $P < 0.05$. **n**, Schematic of collagen patch generation (reconstructed from ref. 26). **o**, Evaluation of mechanical properties of engineered patch, measured by atomic force microscopy. **p, q**, Suture procedure of patch over ischaemic myocardium. **r**, Echocardiography analysis normalized to individual pre-surgery baseline values. **s**, Absolute values of fractional shortening (FS%). **t**, Masson's trichrome staining of the animal cohorts: sham (control, $n = 10$), infarcted mice without treatment (MI only, $n = 8$), MI treated with patch only (MI plus patch, $n = 8$), and infarcted animals treated with patch laden with epicardial conditioned media (MI plus patch plus CM, $n = 8$), 2 weeks after MI. * $P < 0.05$ compared to Sham control, † $P < 0.05$ compared to MI-only, and ‡ $P < 0.05$ compared to MI plus patch (see Methods for details.)

cardiomyocytes (Fig. 2b–d), the transcription levels of myocardial-specific proteins (*Myh6*, *Mlc2v*, and *Mlc2a*, Fig. 2e), and the number of α -actinin⁺ cells with rhythmic Ca²⁺ transients (Fig. 2f). FSTL1 treatment did not induce hypertrophy. Indeed FSTL1 decreased myocyte cell size in a dose-dependent manner (Fig. 2g) in 48 h. Thus, FSTL1 recapitulates the cardiomyogenic activity of the epicardial conditioned medium.

Dynamic expression of Fstl1 after ischaemic injury

During fetal development, endogenous Fstl1 is expressed throughout the myocardium of the primitive heart tube¹², but becomes restricted to the epicardium by mid-gestation (Fig. 2h). Epicardial expression persists throughout adulthood (Fig. 2i–k). Remarkably, Fstl1

localization shifts strikingly following ischaemic injury, such that it becomes abundant in the myocardium (Fig. 2i–l) and absent in the epicardium and infarcted area (Fig. 2i, l).

Epicardial FSTL1 promotes regeneration

Prior studies showed that transient overexpression of Fstl1, either by myocardial transgenic expression (Fstl1-TG¹³, Extended Data Fig. 2a, b) or systemic infusion, is anti-apoptotic following acute ischaemia-reperfusion (I/R)^{14,15}. In the context of permanent myocardial infarction, myocardial Fstl1-TG mice did not recapitulate the effect of the patch containing EMC-media (Extended Data Fig. 2a–j). We assessed epicardial-hFSTL1 delivery by collagen patches loaded with 10 μ g of recombinant bacterial-synthesized hFSTL1 per patch. Patches

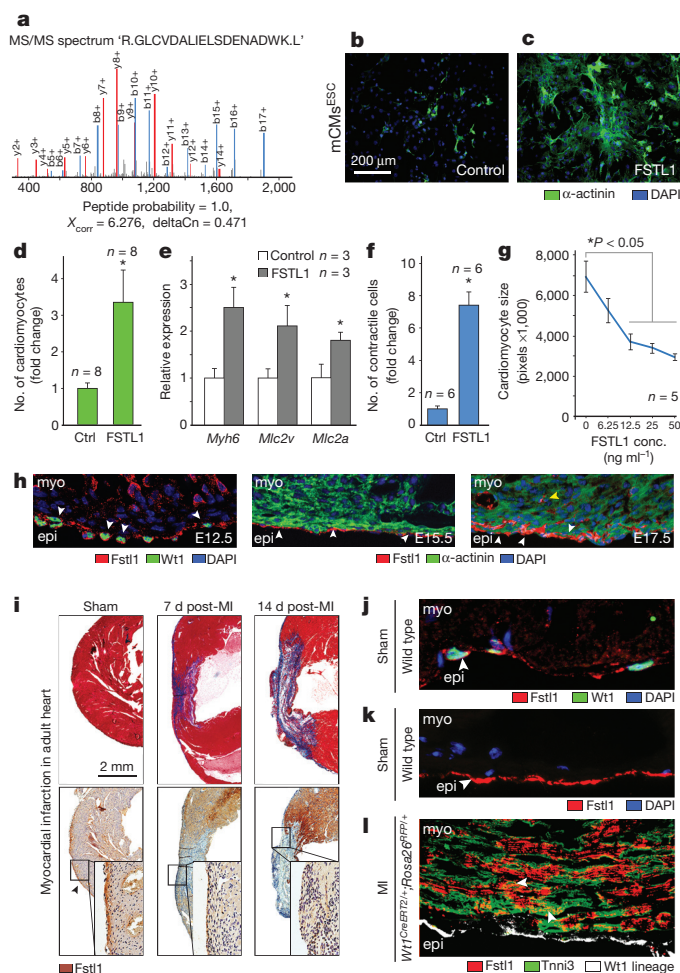


Figure 2 | Fstl1 is an epicardial cardiogenic factor with dynamic expression after ischaemic injury. **a**, MS/MS spectrum of Fstl1. **b–g**, Fstl1 treatment of mCMEs^{ESC} cardiomyocytes measured by immunostaining (α -actinin, green) (**b**, **c**), quantification of myocyte number (**d**), expression of cardiac-specific markers (**e**), cardiomyocytes with rhythmic calcium transient (**f**), and individual cardiomyocyte cell size (**g**). * $P < 0.05$ indicates statistically significantly different from control. **h**, Fstl1 immunostaining in the mouse embryonic heart (days E12.5, E15.5 and E17.5). Fstl1 (red), Wt1 (epicardial marker), α -actinin (myocardial marker), DAPI (nuclei). Fstl1 is expressed in epicardium (white arrowheads), no myocardium (yellow arrowhead). **i**, Expression shift of Fstl1 in the mouse heart after MI. Trichrome staining (upper), labels fibrosis (blue) Fstl1 immunohistochemistry (lower panels, brown). In injured hearts Fstl1 expression is depleted from the epicardium (brown) and upregulated in the myocardium. **j–l**, High resolution images of Fstl1 expression-shift after MI (see Methods for details).

retained immune-detectable hFSTL1 up to 21 days *in vitro*, and 28 days *in vivo*, the longest lengths of times tested (Extended Data Fig. 3a–f). Application of hFSTL1-loaded patches simultaneously with MI significantly improved survival (Fig. 3a) and sustained long-term recovery of cardiac function (Fig. 3b and Extended Data Table 2). Epicardial patch with FSTL1 also improved cardiac function when applied onto infarcted hearts of Fstl1-TG mice (Fig. 3c); thus, myocardial overexpression of FSTL1 is insufficient for long-term recovery but epicardial reconstitution of recombinant FSTL1 is necessary to induce the beneficial effects.

The improved cardiac function and survival following patch with FSTL1 treatment was accompanied by attenuated fibrosis (Fig. 3d, e, Extended Data Fig. 3g–j, Extended Data Table 2, and Supplementary Videos 3–5), increased vascularization of the patch and underlying myocardium at the border of the infarcted region (Fig. 3f–i), as reflected by the increased area occupied by vessels (Fig. 3f, g),

and the increased number of vessels (of any size) per area unit (Fig. 3h, i). Masson's trichrome staining showed contiguous engraftment of the patch with FSTL1 onto the host myocardium and demonstrated migration of host cells into the patch including evidence of striated cells (green arrows, last two columns in Extended Data Fig. 3k).

A similar recovery was found when FSTL1-loaded patches were grafted one week after I/R injury in the mouse, when cardiac function had substantially decreased (about 15% reduction in fractional shortening, FS%). As is typical, cardiac function of untreated animals progressively declined (22%, 20% and 16% fractional shortening at 1, 3 and 5 weeks post-I/R). In contrast, the patch with FSTL1 cohort showed a nearly complete and stable recovery of fractional shortening (to 34% three weeks post-I/R) (Extended Data Fig. 4a–d and Extended Data Table 3), suggesting that epicardial-delivered FSTL1 is sufficient to revert the loss of cardiac function after experimental MI.

FSTL1 induces cardiomyocyte proliferation *in vivo*

Four weeks following MI, the patch with FSTL1 cohort showed evidence of striated myocytes (α -actinin⁺ cells) within the patch (Extended Data Fig. 5a–d). Cardiomyocytes in the border zone had undergone cell division (Extended Data Fig. 5e–h) by several independent criteria, including an increased number of double-positive α -actinin⁺, phospho-histone H3 (pH3)⁺ cells (Fig. 3j–l; and Extended Data Fig. 5e–k), increased incidence of aurora B kinase localized to the midbody between α -actinin⁺ cells (Fig. 3m, n and Supplementary Video 6), and increased incidence of cells that were double-positive for pH3 and the nuclear cardiomyocyte maker PCMI (ref. 16) (Fig. 3o, p) relative to MI and patch-only cohorts (Extended Data Fig. 5l–r). Thus, epicardial FSTL1 delivery activates cardiomyocyte cell cycle entry and cytokinesis *in vivo* reminiscent of the *in vitro* results above. Proliferating cardiomyocytes were found only in the border zone and, to a lesser extent, the infarcted area (Extended Data Fig. 5s, t). Increased cardiomyocyte proliferation was also observed in the I/R injury model with delayed patch implantation (Extended Data Fig. 4e, f). Notably, FSTL1 did not diminish cardiomyocyte apoptosis, the extent of the infarcted area or the area at risk (hypoperfused area) acutely after MI; nor did it affect apoptosis or inflammation at day 4 and day 8 post-MI (Extended Data Fig. 6).

In contrast to patch with FSTL1 delivery, transgenic overexpression of Fstl1 (Fstl1-TG mice) did not show any evidence of cardiomyocyte proliferation after MI (Extended Data Fig. 2k, l), despite increased vascularization described previously¹³ (Extended Data Fig. 2m, n), indicating that epicardial-delivered FSTL1 might function differently than myocardial-expressed Fstl1.

To distinguish whether the FSTL1-responsive cells arise from pre-existing myocytes (*Myh6*⁺ cells) or *de novo* from a progenitor population, we heritably labelled *Myh6*⁺ cardiomyocytes using a 4-OH-tamoxifen-inducible *cre*¹⁷ before injury (Fig. 3q). 4-OH tamoxifen injected into *Myh6*^{mERcre}; *Rosa26*^{Z/EG} mice¹⁷ efficiently labelled pre-existing cardiomyocytes with eGFP before MI (Fig. 3r). Four weeks after patch engraftment, eGFP⁺, pH3⁺ double-positive cells were visible in the infarct area and border zone (Fig. 3s–v), indicating that the proliferating cardiomyocytes expressed *Myh6* before MI. We treated cardiomyocytes at different stages of differentiation with FSTL1 in order to determine which stage(s) can respond by proliferating. Neither adult mouse cardiomyocytes (Extended Data Fig. 7a–f), neonatal rat cardiomyocytes (Extended Data Fig. 7g–j) nor cardiomyogenic progenitor cells (Lin[−], Sca1⁺, SP⁺)¹⁸ responded (Extended Data Fig. 7k–m). Of the cells tested, only immature cardiomyocytes (mCMEs^{ESC}) proliferated in response to FSTL1 (Fig. 4a–f).

It remained paradoxical that neither the endogenous Fstl1 induced by MI nor myocardially overexpressed Fstl1 could induce a regenerative response (Fig. 3 and Extended Data Fig. 2). Western blot analysis indicated that myocardially overexpressed Fstl1 (in neonatal rat ventricular myocytes, NRVC) migrates substantially slower in

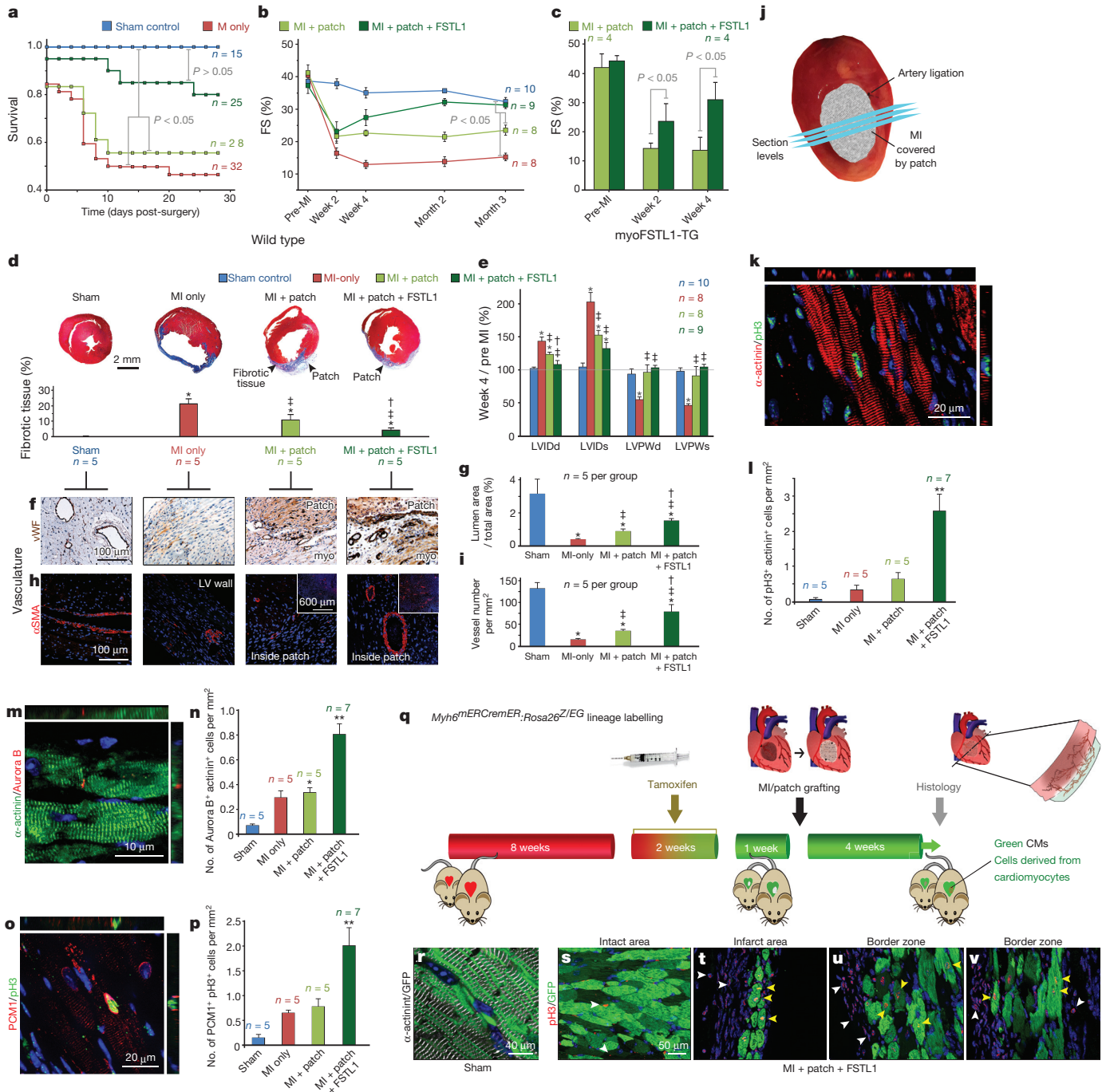


Figure 3 | FSTL1 recapitulates the *in vivo* restorative effect of epicardial conditioned media in the engineered epicardial patch, and promotes cardiomyocyte proliferation. **a**, **b**, Survival (**a**) and kinetics of FS (%) (**b**) analyses after MI in the indicated treatments. **c**, Effect of epicardial hFSTL1 patches on FS% in *Fstl1*-TG mice. **d–i**, Masson's trichrome staining (**d**), morphometric analysis by echocardiography (**e**), and vascularization analysis (**f–i**) 4 weeks after MI. $*P < 0.05$ compared to sham, $\dagger P < 0.05$ vs MI only, and $\ddagger P < 0.05$ vs MI plus patch. **j**, Cross-sections covering infarct/patch area separated 250 μ m, 1–2 mm from apex used for cardiomyocytes proliferation analysis (**k–p**), 4 weeks after MI. **k**, **m**, **o**, Co-staining of pH3 and α -actinin (**k**), midbody-localized aurora B kinase between α -actinin⁺ cells

(**m**), and double-positive cells for pH3 and PCM1 (cardiomyocyte nuclei¹⁶) (**o**) 4 weeks post-MI, quantified in **l**, **n**, **p**, normalized to myocardium area quantified by trichrome staining of immediately adjacent section. $*P < 0.05$ from sham. $**P < 0.05$ from all other groups. **q–v**, Lineage tracing of FSTL1-responsive cells in 4-OH-tamoxifen treated *Myh6^{mERCreMER}; Rosa26^{Z/EG}* mice; patch with FSTL1 applied simultaneously to MI, and hearts were collected 4 weeks post-MI (**q**) with efficient labelling of cardiomyocytes (**r**). Infarcted hearts showing eGFP⁺ (pre-existing, green) cardiomyocytes positive for pH3 (yellow arrowheads) (white arrowheads: pH3⁺ eGFP⁻ cells) (**s–v**). (length of treatment (**a–c**) and n for each experiment indicated in graph, see Methods for details).

SDS-polyacrylamide gels than does epicardially synthesized *Fstl1* (EMC), and that tunicamycin treatment eliminates the difference (Fig. 4g), suggesting cell-type specific glycosylation. The bacterially produced recombinant human FSTL1 (as used in the patch) showed a faster migration consistent with less extensive glycosylation (Fig. 4h,

i). Direct comparison of recombinant human FSTL1 produced in bacterial versus mammalian cells (NS0-derived mouse myeloma cell line) revealed that mammalian-expressed FSTL1, but not bacterial FSTL1, protects mCMs^{ESC} from H₂O₂-induced apoptosis (Fig. 4j), consistent with evidence of cardioprotection¹⁴ but not regeneration

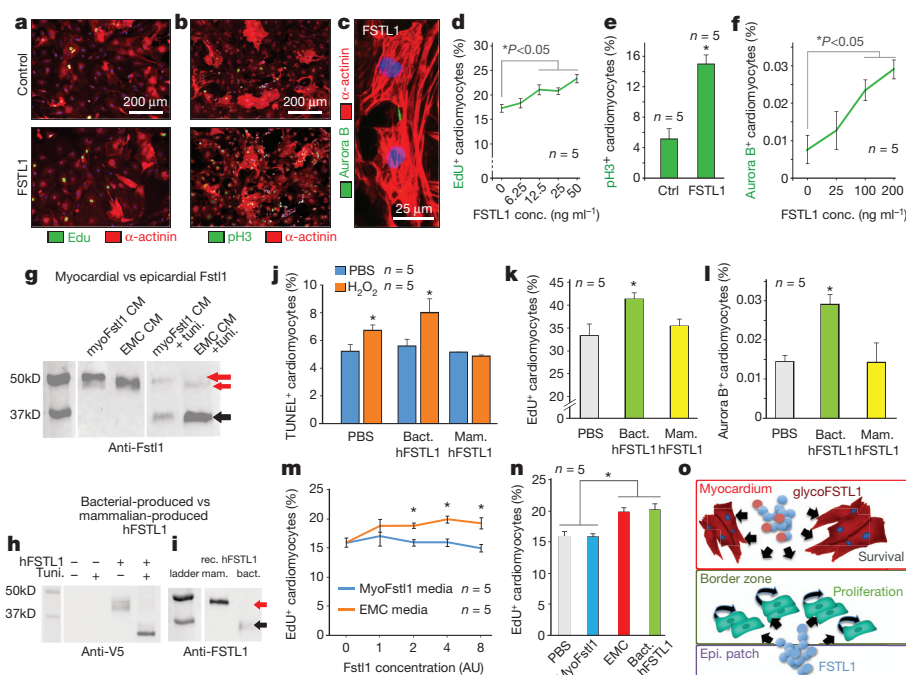


Figure 4 | FSTL1 proliferative activity on early cardiomyocytes depends on the cells' selective post-transcriptional FSTL1 modifications. **a–f**, FSTL1 promotes proliferation of mCMs^{ESC}, measured by EdU incorporation (**a**), pH3 (**b**), and aurora B immunostaining (**c**), and quantified in **d–f**. **g–i**, Western blot analysis of Fstl1 secreted in cultured cardiomyocytes (myoFSTL1 CM) infected with Adeno-Fstl1 and in EMC (EMC CM) in the presence of tunicamycin (glycosylation inhibitor) (**g**), hFSTL1-V5 tagged expressed in AD-293 cells (**h**), and mammalian and bacterial-produced FSTL1 (**i**). Red arrows, glycosylated; black arrows, hypoglycosylated. **j**, Mammalian-produced FSTL1

attenuates H₂O₂ induced apoptosis, while bacterial-produced FSTL1 cannot. **k, l**, Bacterially-produced FSTL1 promotes mCMs^{ESC} EdU incorporation and aurora B positivity whereas mammalian-produced FSTL1 does not. **m, n**, Quantification of EdU incorporation in mCMs^{ESC} treated with conditioned media of EMC and Fstl1-overexpressing NRVC (concentration normalized to Fstl1 content). **P* < 0.05 indicates statistically different from control (see Methods for details). **o**, Working model of FSTL1 in distinct cardiac compartments.

(Extended Data Fig. 2a–j) in Fstl1 TG mice. In contrast, bacterially synthesized human FSTL1 promotes mCMs^{ESC} proliferation, whereas human FSTL1 produced in NS0-derived cells or in NRVCs cannot stimulate proliferation of mCMs^{ESC} (Fig. 4k–n). Thus, whether FSTL1 induces cardioprotection versus proliferation correlates with cell source and might reflect post-translational modification.

Epicardial FSTL1 in a preclinical swine model

Epicardial delivery of FSTL1 was evaluated in the swine model of I/R injury. I/R decreased left ventricular ejection fraction (EF %) from ~50% before MI, as determined by magnetic resonance imaging (MRI), to ~30% at 1 week after injury. Application of the patch with FSTL1 to the epicardium over the injured tissue at this time (1 week post-MI I/R) stimulated recovery of contractile function (to ~40% EF) in 2 weeks (3 weeks post-MI I/R) (Fig. 5a, b). The recovery remained stable for an additional 2 weeks, the longest time analysed, and was in contrast to the steady decline seen without treatment or following treatment with patch alone (Fig. 5b). FSTL1-treated pigs demonstrated the least scar size of all treatments, including the patch-only condition (see representative MRI images (Fig. 5c, d)). Examination of histological sections of tissues 4 weeks after patch implantation confirmed the limited fibrosis and showed integration of the patch into the host tissue (Fig. 5e). Cardiomyocytes in the border zone and ischaemic area of the patch with FSTL1 treated hearts also had evident EdU labelling (Fig. 5i–m) and midbody-localized aurora B kinase (indicative of cytokinesis) (Fig. 5n). Vascular smooth muscle cells were also EdU⁺ suggestive of arteriogenesis (Fig. 5g, h). Thus, the patch with FSTL1 appears therapeutic in the swine MI I/R model.

Discussion

Heart regeneration studies in zebrafish suggested that the epicardium is activated by injury to produce factors and cells that sustain cardiac

function¹⁹. Unlike lower vertebrate hearts, which are robustly regenerative, the mammalian heart retains negligible regenerative potency in adulthood and, instead, sustains cardiomyocyte death and scarring following injury. Very little is known of the endogenous mechanisms that limit regeneration and the topic remains a subject of intense therapeutic interest and scientific debate²⁰. Our data suggest a new view of epicardial function after injury in the mammalian heart. Rather than activation to support cardiac function, the loss of epicardial FSTL1 expression after injury, and the functional and anatomical recovery by reconstitution in an engineered biomaterial, indicate that ischaemic injury induces a maladaptive loss of FSTL1 in the epicardium.

We sought to identify the cell population that proliferates in response to FSTL1. FSTL1 could not stimulate mature adult ventricular cardiomyocytes to synthesize DNA or divide, nor did it induce hypertrophy (as can occur in response to mitogens) either at 48 h (Fig. 2g) or 4 weeks post-MI (Extended Data Fig. 5o). In contrast, FSTL1 stimulated replication of newly emerging cardiomyocytes from mouse ESC cultures (Fig. 4a–f). FSTL1 did not enhance replication of either ESC-derived progenitors before the appearance of *Myh6* (not shown) or a population of cardiac progenitors isolated from the adult murine heart (Extended Data Fig. 7k–m), suggesting that competence to respond to FSTL1 occurs transiently. Although at least some mononuclear adult cardiomyocytes can be induced to divide²¹, the FSTL1-responsive cardiomyocytes in our experiments have even less mature sarcomeric and electrophysiological properties²² (for example, automaticity, relatively high maximum diastolic potential, and slow action potential peak *V*_{max}) (Extended Data Fig. 1). The cells that respond to FSTL1 might overlap cells identified in an earlier analysis of infarcted hearts labelled with *Myh6-cre*, in which a minor population of *cre*-labelled cells were reported to divide and give rise new cardiomyocytes upon ischaemic injury²³. Whether the FSTL1-responsive cells

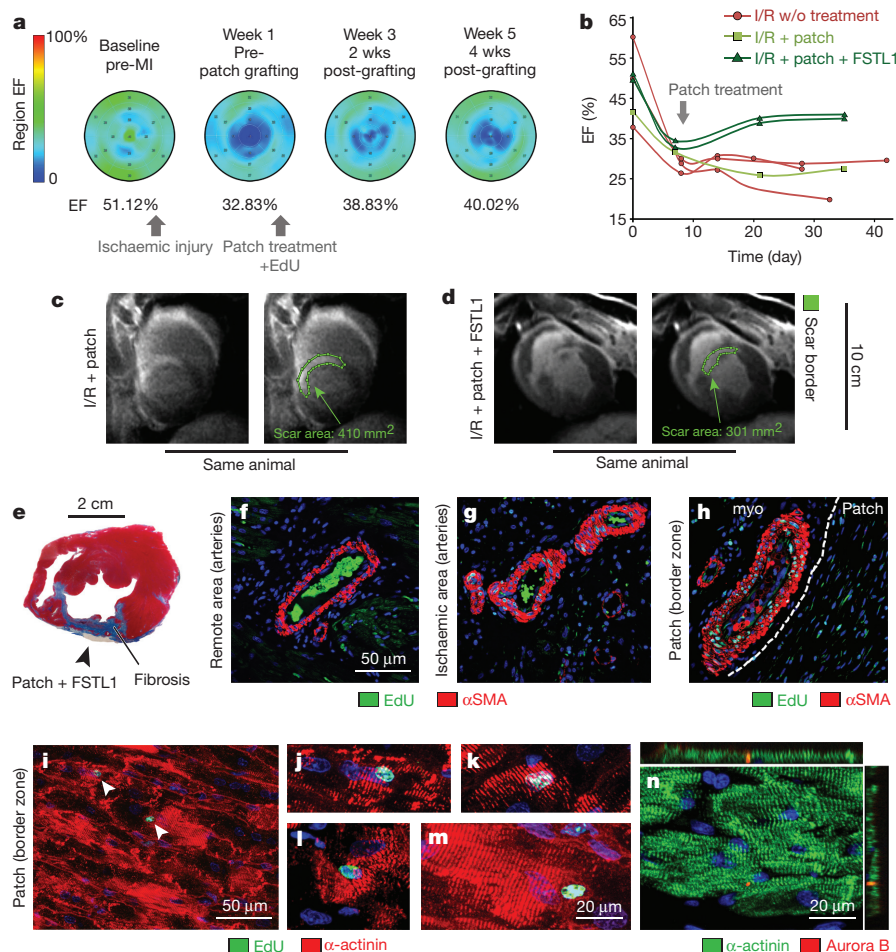


Figure 5 | Epicardial FSTL1 delivery activates cardiac regeneration in preclinical model of ischaemic heart injury. **a–d**, Time course MRI analysis of cardiac function in pigs. Functional analysis by measurement of ejection fraction (EF%) (**a**, **b**). Scar size at week 4 post-grafting (**c**, **d**). Green lines highlight scar perimeter. **e–n**, Analysis at week 4 post-grafting. Masson's

reflect resident *Myh6*⁺ recruited upon injury for example²³, or derive from de-differentiation²⁴ (thus recapitulating the zebrafish model²⁵) is an interesting question whose resolution will depend on improved method to identify and/or isolate such cells.

Myocardial Fstl1 induced by MI cannot promote a regenerative response, either basally or when abundantly overexpressed transgenically in cardiomyocytes (Extended Data Fig. 2). However, transgenic myocardial Fstl1 is cardioprotective post-MI¹⁴. Direct comparison of FSTL1 overexpressed in cardiomyocytes versus the epicardial (EMC) protein revealed tunicamycin-sensitive differences in SDS-PAGE mobility (Fig. 4), consistent with the possibility of differential glycosylation (or other post-translational modification) depending on the cell in which it is expressed. We infer from these data that native epicardial and myocardial FSTL1 have analogous differences in glycan structure that affect their function. It will be important to determine the structure of the glycans, as well as elucidate how post-translation modifications dictate whether FSTL1 promotes anti-apoptosis (myocardial) or cardiomyocyte proliferation (EMC and bacterially produced).

These studies identified FSTL1 as a regenerative factor that is normally present in healthy epicardium, but lost upon MI, suggesting a mechanism whereby injury maladaptively diminishes the regenerative potency of the mammalian heart. Reconstitution of FSTL1 by an engineered epicardial biomaterial improved cardiac function in the mouse MI, mouse MI I/R and preclinical swine MI I/R models with

trichrome staining (**e**). EdU incorporation (newly synthesized DNA) in the vascular smooth muscle cells (**f–h**). White line demarcates patch and host tissue. **i–n**, EdU (**i–m**) incorporation and aurora B kinase positivity (**n**) in cardiomyocytes at week-4 post-grafting (see Methods for details).

evidence of cardiomyocyte regeneration amenable to clinical translation.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 18 December 2013; accepted 11 August 2015.

Published online 16 September; corrected online 23 September 2015

(see full-text HTML version for details).

- van Wijk, B., Gunst, Q. D., Moorman, A. F. & van den Hoff, M. J. Cardiac regeneration from activated epicardium. *PLoS ONE* **7**, e44692 (2012).
- Cai, C. L. *et al.* A myocardial lineage derives from Tbx18 epicardial cells. *Nature* **454**, 104–108 (2008).
- Lavine, K. J. & Ornitz, D. M. Rebuilding the coronary vasculature: hedgehog as a new candidate for pharmacologic revascularization. *Trends Cardiovasc. Med.* **17**, 77–83 (2007).
- Brade, T. *et al.* Retinoic acid stimulates myocardial expansion by induction of hepatic erythropoietin which activates epicardial Igf2. *Development* **138**, 139–148 (2011).
- Mellgren, A. M. *et al.* Platelet-derived growth factor receptor beta signaling is required for efficient epicardial cell migration and development of two distinct coronary vascular smooth muscle cell populations. *Circ. Res.* **103**, 1393–1401 (2008).
- Smart, N. *et al.* Myocardial regeneration: expanding the repertoire of thymosin beta4 in the ischemic heart. *Ann. NY Acad. Sci.* **1269**, 92–101 (2012).
- Kikuchi, K. *et al.* *tcf21* + epicardial cells adopt non-myocardial fates during zebrafish heart development and regeneration. *Development* **138**, 2895–2902 (2011).
- Zhou, B. *et al.* Adult mouse epicardium modulates myocardial injury by secreting paracrine factors. *J. Clin. Invest.* **121**, 1894–1904 (2011).

9. Serpooshan, V. *et al.* The effect of bioengineered acellular collagen patch on cardiac remodeling and ventricular function post myocardial infarction. *Biomaterials* **34**, 9048–9055 (2013).
10. Tanaka, M. *et al.* DIP2 disco-interacting protein 2 homolog A (*Drosophila*) is a candidate receptor for follistatin-related protein/follistatin-like 1—analysis of their binding with TGF- β superfamily proteins. *FEBS J.* **277**, 4278–4289 (2010).
11. Widera, C. *et al.* Identification of Follistatin-Like 1 by expression cloning as an activator of the growth differentiation factor 15 gene and a prognostic biomarker in acute coronary syndrome. *Clin. Chem.* **58**, 1233–1241 (2012).
12. Adams, D., Larman, B. & Oxburgh, L. Developmental expression of mouse Follistatin-like 1 (Fstl1): Dynamic regulation during organogenesis of the kidney and lung. *Gene Expr. Patterns* **7**, 491–500 (2007).
13. Shimano, M. *et al.* Cardiac myocyte follistatin-like 1 functions to attenuate hypertrophy following pressure overload. *Proc. Natl Acad. Sci. USA* **108**, E899–E906 (2011).
14. Oshima, Y. *et al.* Follistatin-Like 1 is an Akt-regulated cardioprotective factor that is secreted by the heart. *Circulation* **117**, 3099–3108 (2008).
15. Ogura, Y. *et al.* Therapeutic impact of follistatin-like 1 on myocardial ischemic injury in preclinical models. *Circulation* **126**, 1728–1738 (2012).
16. Bergmann, O. *et al.* Identification of cardiomyocyte nuclei and assessment of ploidy for the analysis of cell turnover. *Exp. Cell Res.* **317**, 188–194 (2011).
17. Sohal, D. S. *et al.* Temporally regulated and tissue-specific gene manipulations in the adult and embryonic heart using a tamoxifen-inducible Cre protein. *Circ. Res.* **89**, 20–25 (2001).
18. Oh, H. *et al.* Cardiac progenitor cells from adult myocardium: homing, differentiation, and fusion after infarction. *Proc. Natl Acad. Sci. USA* **100**, 12313–12318 (2003).
19. Lepilina, A. *et al.* A dynamic epicardial injury response supports progenitor cell activity during zebrafish heart regeneration. *Cell* **127**, 607–619 (2006).
20. Mercola, M., Ruiz-Lozano, P. & Schneider, M. D. Cardiac muscle regeneration: lessons from development. *Genes Dev.* **25**, 299–309 (2011).
21. Bersell, K., Arab, S., Haring, B. & Kuhn, B. Neuregulin1/ErbB4 signaling induces cardiomyocyte proliferation and repair of heart injury. *Cell* **138**, 257–270 (2009).
22. Chen, H. S., Kim, C. & Mercola, M. Electrophysiological challenges of cell-based myocardial repair. *Circulation* **120**, 2496–2508 (2010).
23. Senyo, S. E. *et al.* Mammalian heart renewal by pre-existing cardiomyocytes. *Nature* **493**, 433–436 (2013).
24. Zhang, Y. *et al.* Dedifferentiation and proliferation of mammalian cardiomyocytes. *PLoS ONE* **5**, e12559 (2010).
25. Jopling, C. *et al.* Zebrafish heart regeneration occurs by cardiomyocyte dedifferentiation and proliferation. *Nature* **464**, 606–609 (2010).
26. Brown, R. A., Wiseman, M., Chuo, C. B., Cheema, U. & Nazhat, S. N. Ultrarapid engineering of biomimetic materials and tissues: fabrication of nano- and microstructures by plastic compression. *Adv. Funct. Mater.* **15**, 1762–1770 (2005).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank L. M. Brill for spectrometry, F. Cerignolli for providing valuable cells, S. Metzler and P. Kim for help with imaging, and P. Shah for assistance in mouse and pig experiments. This work was supported by NIH grants to P.R.-L. (HL065484 and R01 HL086879); M.Me. (HL113601, HL108176, P01 HL098053); P30 AR061303 and P30 CA030199 for shared services and by the California Institute for Regenerative Medicine (CIRM, RC1-00132) to M.Me. K.We. and C.H. were SBMRI CIRM postdoctoral fellows (TG2-0116). V.S. was an Oak Foundation postdoctoral fellow. Support was also provided by NIH/NHLBI 5UM1 HL113456 to P.C.Y.; HL116591 to K.Wa., M.J.B. was supported by the NIH (K08 AI079268) and the Stanford BioX Interdisciplinary Initiatives Program and NSF NSEC(PHY-0830228). B.Z. was supported by the National Basic Research Program of China (2013CB945302 and 2012CB945102) and the National Natural Science Foundation of China (91339104, 31271552, and 31222038). A Seed Grant to P.R.-L. from the Stanford Cardiovascular Institute supported the swine study.

Author Contributions K.We. and W.C. performed experiments on EMCs and mESCs. K.We. generated mCMs^{ESC}, and performed cardiomyogenic and proliferation assays on mCMs^{ESC}, proliferation assays on NRVC. K.We. and W.Z. performed mass spectrometry experiments. M.D.-C. performed immunostaining of Fstl1. K.We. and A.S. performed calcium transient experiment. V.S. generated cardiac patch. V.S., A.W., and M.J.B. performed biomechanical analysis of cardiac patch biomaterial. M.Z. and V.S. performed mouse MI experiments and echocardiography. Y.M. and P.C.Y. performed MRI analysis. K.We. and W.C. performed immunostaining of α -actinin, pH3 and aurora B. K.We. performed GFP, TUNEL and f4/80 staining. K.We. and P.B. performed analysis of high throughput imaging results. M.Mh. analysed the release of Fstl1 from the patch *in vitro*. V.S. and G.F. performed assays of adult mouse cardiomyocytes. K.Wa., K.N. and S.M. performed experiments with Fstl1-TG mice. B.Z. performed experiments with adult mouse epicardium-conditioned media. X.T., Q.L. and B.Z. performed *Wt1*^{CreERT2} lineage tracing studies and experiments with adult mouse epicardium conditioned medium. K.Wa. and S.M. provided data on systemic delivery of FSTL1. M.N. and M.D.S. provided data on myocyte progenitors. M.J.B. supervised AFM studies. D.B. supervised and coordinated *in vivo* mouse physiology experiments. V.S., Y.M., and P.C.Y. conducted the preclinical swine study. C.H. performed FSTL1 overexpression and western blot experiments. K.We., V.S., M.Me. and P.R.-L. designed experiments and prepared the manuscript. M.J.B.v.d.H. provided materials.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to P.R.-L. (prlozano@stanford.edu).

METHODS

Data reporting. No statistical methods were used to predetermine sample size.

Cell preparation.

Progenitor cells $Sca1^{+}, Myh6^{-}$ cardiomyocyte progenitors were obtained by the Schneider laboratory as described¹⁸ (Extended Data Fig. 7k–m).

Epicardial mesothelial cells (EMCs) These were maintained in DMEM with 10% FBS and antibiotics/antimycotic as described²⁷. EMCs are stably transduced with H2B–mCherry lentivirus for nuclei labelling (Figs 1 and 4).

Mouse embryonic stem cell-derived cardiomyocytes (mCMEs) A stable mouse ESC line for drug resistance selection of cardiomyocytes (*Myh6-Puro^r;Rex-Blast^r*) was generated by lentiviral transduction and blasticidin selection, similarly to our previously reported human line²⁸.

mCMEs These were obtained by differentiation of *Myh6-Puro^r;Rex-Blast^r* mESCs in a differentiation media containing: Iscove's Modified Dulbecco Media (IMDM) supplemented with 10% FBS, 2 mM glutamine, 4.5×10^{-4} M monothioglycerol, 0.5 mM ascorbic acid, $200 \mu\text{g ml}^{-1}$ transferrin (Roche), 5% protein-free hybridoma media (PFHM-II, Invitrogen) and antibiotics/antimycotic as embryoid bodies (EBs) until day 4 and plated onto adherent cell culture plate until 9, one day after the onset of spontaneous beating. To purify *Myh6⁺* cardiomyocytes, puromycin was added at differentiation day 9 for 24 h. Subsequently cells were trypsinized and plated as monolayer cardiomyocytes. Conditioned media and FSTL1 treatments were typically performed 24 h after monolayer plating. The length of the treatments is indicated in each figure legends (Figs 1, 2, 4 and Extended Data Fig. 1).

AD-293 These cells were directly purchased from Stratagene avoiding misidentification, and cultured in DMEM media with 10% FBS and with pen/strep. It's used for its high transfection efficiency and yield of recombinant proteins (Fig. 4h).

EMCs and *Myh6-Puro^r;Rex-Blast^r* mESCs, and AD-293 cells are quarterly tested for mycoplasma contamination when in use.

Embryonic cardiomyocytes We used fluorescence activated cell sorting (FACS) to purify cardiomyocytes from *Tnt-cre;Rosa26^{mTmG/+}* (C57BL/6J) and ICR mixed background) hearts from e12.5 embryos. Hearts were dissociated collagenase IV digestion and GFP⁺ cells for FACS purification. The GFP⁺ cells were cultured and confirmed to be cardiomyocytes by their expression of the cardiomyocyte specific markers alpha actinin (ACTN2) and cardiac troponin T (TNNT2). They were rhythmically beating when cultured *in vitro* (Fig. 1j–m).

Neonatal rat ventricular cardiomyocytes (NRVCs) These cells were isolated with the neonatal rat cardiomyocyte isolation kit (Cellutron) and cultured at 37 °C with 5% CO₂. In brief, ventricles were dissected from 1–2-day-old Hsd:s.d. rats (Sprague Dawley), then digested five times for 15 min each with the enzyme cocktail at 37 °C. Cells were pooled, pre-plated for 90 min on an uncoated cell culture dish to remove fibroblasts, and plated on 1% gelatin-coated cell culture plastic dishes in high-serum media (DME/F12 [1:1], 0.2% BSA, 3 mM sodium-pyruvate, 0.1 mM ascorbic acid, 4 mg l^{-1} transferrin, 2 mM L-glutamine, and 5 mg l^{-1} ciprofloxacin supplemented with 10% horse serum and 5% fetal bovine serum (FBS)) at 3×10^5 cells per cm². After 24 h, media was changed to low-serum medium (same but with 0.25% FCS) and cells cultured until use (Fig. 4g, m, n and Extended Data Fig. 7g–j).

Adult mouse cardiomyocytes These were isolated from 3 month old *Myh6^{mERcre};Rosa26^{Z/EG}* C57BL/6J mice as previously published²⁹. Briefly, mice were anesthetized with pentobarbital sodium (100 mg per kg IP). The heart was removed and retrograde perfused at 37 °C with a Ca²⁺-free solution (in mM, 120 NaCl, 14.7 KCl, 0.6 KH₂PO₄, 0.6 Na₂HPO₄, 1.2 MgSO₄·7H₂O, 4.6 NaHCO₃, 10 Na-HEPES, 30 taurine, 10 BDM, 5.5 glucose) followed by enzymatic digestion with collagenase. Ventricles were cut into small pieces and further digested. Stop buffer (Ca²⁺-free solution, 12.5 μM CaCl₂, 10% bovine calf serum) was added and the cell suspension was centrifuged at 40 g for 3 min. Myocytes were resuspended in stop buffer in increasing CaCl₂ concentrations until 1 mM was achieved. Cells were then resuspended in MEM, 5% bovine calf serum, 10 mM BDM, 2 mM L-glutamine and added to the collagen solution, pre-polymerization (250,000 cells per ml or per patch). Following collagen gelation and plastic compression, cellular patches were cultured in aforementioned (plating) media overnight and then transferred into culture media: MEM, 1 mg ml⁻¹ bovine serum albumin, 25 μM blebbistatin, 2 mM L-glutamine, in presence or absence of recombinant FSTL1 (AVISERA BIOSCIENCE, 10 ng ml⁻¹). At day 7, fluorescent ubiquitination-based cell-cycle indicator (FUCCI, Premo FUCCI Cell Cycle Sensor, Life Technologies, US) assay was conducted on the 3D culture specimens as previously described³⁰. In this assay, G1 and S/G2/M cells emit red and green fluorescence, respectively. The volume of

Premo geminin-GFP and Premo Cdt1-RFP reagent were calculated using the equation below:

$$\text{Volume (ml)} = \frac{\text{number of cells} \times \text{PPC}}{1 \times 10^8}$$

where the number of cells is the estimated total number of cells at the time of cell labelling (equal to CM seeding density, PPC (particles per cell) is the number of viral particles per cell (40 in this assay), and 1×10^8 is the number of viral particles per ml of the reagent. The volumes of reagents calculated above were directly added to the cellular patches in complete cell medium, mixed gently, and incubated overnight in the culture incubator (≥ 16 h). Patch samples were imaged using a conventional fluorescence microscope, using GFP and RFP filter sets (Extended Data Fig. 7a–f).

Co-culture experiments. mCMEs are co-cultured with H2B–mCherry EMCs for 4 days and visualized by α -actinin immunofluorescent staining and H2B–mCherry fluorescence (Fig. 1a, b), cardiomyocyte counting (Fig. 1c, $n = 3$), and cardiogenic gene expression normalized to *Gapdh* gene expression (Fig. 1d $n = 3$).

Epicardial conditioned media. Rat epicardial mesothelial cells (EMC) conditioned media EMC²⁷ cells were cultured in 10% FBS DMEM with penicillin/streptomycin until confluent ($\sim 1 \times 10^6 \text{ cm}^{-2}$), then washed with PBS three times and media is changed to serum free DMEM with penicillin/streptomycin without phenol red and cultured for 2 additional days before the media was collected as conditioned media (20 ml of media is added for conditioning and 18 ml is collected after 2 days). Collected media was filtered through 0.22 μm pore membrane (Millipore). Control conditioned media were prepared the same way but without EMC cells (Fig. 1e–i, n–t).

Neonatal rat ventricular cardiomyocytes (NRVCs) conditioned media. NRVC were infected with adenovirus expressing un-tagged mouse Fstl1 at MOI 50. 24 h post-infection culture media was replaced by serum free media (DMEM/F12 with penicillin/streptomycin). The media was conditioned with the infected NRVC and EMC cells for 24 h (Fig. 4m, n).

mCMEs were treated with control and EMC-conditioned media for 8 days before α -actinin immunofluorescent staining (Fig. 1e, f), cardiomyocyte counting (Fig. 1g, $n = 3$), analysis of cardiogenic gene expression normalized to *Gapdh* gene expression (Fig. 1h, $n = 3$) and quantification of the number of cardiomyocytes with rhythmic calcium transient measured automatically using a Kinetic Imaging Cytometer (Vala Sciences) (Fig. 1i, $n = 3$).

mCMEs were treated with serial dilutions of conditioned media of EMC and Fstl1-overexpressing NRVC for 24 h with $10 \mu\text{g ml}^{-1}$ EdU, and stained for α -actinin and EdU (Fig. 4m, n, $n = 5$). The concentrations of the conditioned media are normalized to amount of Fstl1 expression by western blot.

Adult mouse EPDC conditioned media This was generated in the Zhou laboratory⁸. Briefly, eight-week old adult *Wt1^{creERT2/+};Rosa26^{mTmG/+}* hearts mice in C57BL/6J and ICR mixed background were injected orally with 4 mg tamoxifen by gavage, four to five oral injections were administered during a two-week period. Myocardial infarction was then induced by ligation of left anterior descending coronary artery on (11 weeks old) adult mice. One week after injury, we collected *Wt1^{creERT2/+};Rosa26^{mTmG/+}* hearts, which were then digested with collagenase IV into single cells. Digestion solution was made by adding 4 ml 1% collagenase IV and 1 ml 2.5% trypsin into 44.5 ml Hanks' balanced salt solution, and supplemented with 0.5 ml chicken serum and 0.5 ml horse serum. Cells were resuspended in Hank's balanced salt solution, 4 ml digestion solution was added to each tube and rocked gently in 37 °C shaker for 6 min. After removing the supernatant containing dissociated cells, we added another 4 ml digestion solution to repeat the digestion 6 times. After final digestion, we filtrated the cells through 70 μm filter and pellet cells by centrifuging at 200 g for 5 min at 4 °C. Cells were then re-suspended by Hanks' balanced salt solution for FACS isolation. Dissociated cells from GFP- hearts were used as a control for gate setting in FACS. GFP⁺ cells (epicardium-derived cells, EPDCs) were isolated from GFP⁺ *Wt1^{creERT2/+};Rosa26^{mTmG/+}* hearts by FACS and these GFP⁺ purified populations were confirmed to be GFP⁺ cells under fluorescence microscope (Fig. 1j). FSTL1 expression (determined by PCR) was restored in cultured GFP⁺ EPDCs. Complete conditioned media from EPDCs was then added to embryonic cardiomyocytes culture for 48 h before assay for proliferation (Fig. 1k–m, $n = 5$).

MTT Assays. Proliferation of cardiomyocytes treated with conditional medium was measured by MTT assay using Celltiter 96 Aqueous One solution (Promega) as previously described⁸. After adding the Celltiter 96 Aqueous One reagent into the cell culture medium, we incubate the plate at 37 °C for 3–4 h, and then record the absorbance at 490 nm using a 96-well plate reader. Absorbance at 490 nm is tightly correlated with cell number. The MTT readout on the y-axis, labelled MTT assay (A490), thus reflects the relative number of cells from each well between groups of treatment (Fig. 1k). Boiling of conditioned media abolished the

growth-promoting effects (Fig. 1k), suggesting a proteinaceous nature of the effective components.

Recombinant FSTL1 was purchased from AVISCERA BIOSCIENCE (00347-02-100, produced in *E. coli*) and R&D system (1694-FN-050, produced in mouse myeloma cell line, NS0derived).

mCME^{ESC} were treated with bacteria-synthesized recombinant human FSTL1 (10 ng ml⁻¹) for 8 days with media change every 2 days, before α -actinin immunofluorescent staining (Fig. 2b, c), cardiomyocyte counting (Fig. 2d, $n = 8$), analysis of cardiogenic gene expression normalized to *Gapdh* gene expression (Fig. 2e, $n = 3$) and quantification of the number of cardiomyocytes with rhythmic calcium transient measured automatically using a Kinetic Imaging Cytometer (Vala Sciences) (Fig. 2f, $n = 6$). mCME^{ESC} were treated with bacteria-synthesized recombinant human FSTL1 (6.25–50 ng ml⁻¹) for 2 days before measurement of individual cardiomyocyte cell size (in pixels) (Fig. 2g, $n = 5$).

mCME^{ESC} were stimulated with 6.25, 12.5, 25 and 50 ng ml⁻¹ of bacteria produced FSTL1 for 24 h with 10 μ g ml⁻¹ EdU, and stained for α -actinin and EdU (Fig. 4a, d, $n = 5$). mCME^{ESC} were stimulated with 10 ng ml⁻¹ FSTL1 for 48 h and stained for α -actinin and pH3 (Fig. 4b, e, $n = 5$). mCME^{ESC} were stimulated with 25, 100, 200 ng ml⁻¹ FSTL1 for 48 h, and stained for α -actinin (red) and aurora B (Fig. 4c, f, $n = 5$).

mCME^{ESC} were stimulated with 10 nM H₂O₂, and 10 ng ml⁻¹ bacteria and mammalian produced FSTL1 for 24 h, and staining for α -actinin and TUNEL for cell death (Fig. 4j, $n = 5$).

mCME^{ESC} were stimulated with 10 ng ml⁻¹ of bacteria and mammalian produced FSTL1 for 24 h with 10 μ g ml⁻¹ EdU, and stained for α -actinin and EdU (Fig. 4k, $n = 5$), and α -actinin and aurora B (Fig. 4l, $n = 5$).

FSTL1 overexpression and western blot. AD-293 cells were transiently transfected with human FSTL1 plasmid (GE Dharmacon, ID: ccsbBroad304_02639 pLX304-Blast-V5-FSTL1) using lipofectamine 2000 (mocked transfection was done with lipofectamine and no plasmid). 48 h post-transfection serum containing media was replaced by serum free DMEM and incubated with the cells for 24 h. Tunicamycin was used at 2 μ g ml⁻¹. Conditioned media from tunicamycin samples was collected during 16 h (cells looked healthy). Conditioned media was spun at 400 g for 7 min and then concentrated approximately 20 times using Microcon-10 kDa cut off columns (Millipore). Samples were combined 1 to 1 ratio with 2 \times SDS sample buffer containing protease inhibitor, DTT and 5 mM EDTA, boiled for 10 min at 95 °C and run in a 4–15% acrylamide Mini-Protein TGX gel, transferred to nitrocellulose membrane and incubated with anti-V5 primary antibody MAB 15253 (Pierce) 1:1,000 dilution and anti-mouse 800 nm conjugated secondary antibody at 1:10,000 dilution (Odyssey), and scanned using the Odyssey Clx Imager (Fig. 4h).

Neonatal rat ventricular cardiomyocytes were infected with adenovirus expressing un-tagged mouse Fstl1 at MOI 50. 24 h post-infection culture media was replaced by serum-free media. Serum free DMEM/F12 pen/strep media was conditioned with the infected NRVC and EMC cells for 24 h. Tunicamycin was used at 1 μ g ml⁻¹ and media was conditioned for 16 h. Conditioned media was spun at 400 g 7 min and then concentrated using Microcon-10 kDa cut off columns (Millipore). Samples were combined 1 to 1 ratio with 2 \times SDS sample buffer containing protease inhibitor, DTT and 5 mM EDTA, boiled 10 min at 95 °C and run in Any KD Mini-Protein TGX gel, transferred to nitrocellulose membrane and incubated with anti-FSTL1 MAB1694 (R&D) primary antibody 1:500 dilution and anti-rat 800 nm conjugated secondary antibody at 1:10,000 dilution (Odyssey), and scanned using Odyssey Clx Imager. Blocking and antibody incubation was done in Odyssey blocker. The western blot for recombinant FSTL1 (100 ng each) was performed the same way (Fig. 4g, i).

RNA extraction and qRT-PCR. Total RNA was extracted with TRIzol (Invitrogen) and reverse transcribed to cDNA with QuantiTect Reverse Transcription Kit (Qiagen) according to the manufacturer's instructions. cDNA samples synthesized from 100 ng of total RNA were subjected to RT-qPCR with LightCycler 480 SYBR Green I Master kit (Roche) performed with LightCycler 480 Real-Time PCR System (Roche) (Figs 1d, h and 2e and Extended Data Fig. 7b–d). Primer sequences are listed in Supplementary Table 1.

LC-MS/MS analysis of conditioned-media. First, Tris(2-carboxyethyl)phosphine (TCEP) was added into 1 ml of conditional media to 10 mM and the protein sample was reduced at 37 °C for 30 min. Then iodoacetamide was added to 20 mM and the solution was alkylated at 37 °C for 40 min in the dark. Mass spectrometry grade of trypsin (Promega) was then added to the solution as 1:100 ratio. After overnight digestion at 37 °C, the sample was then desalted using a SepPack cartridge, dried using a SpeedVac and re-suspended in 100 μ l of 5% formic acid. The resulting peptides were analysed on-line by an LC-MS/MS system, which consisted of a Michrom HPLC, a 15 cm Michrom Magic C18 column, a low flow ADVANCED Michrom MS source, and a LTQ-Orbitrap XL (Thermo Scientific, Waltham, MA). A 120-min gradient of 0–30% B (0.1%

formic acid, 100% acetonitrile) was used to separate the peptides, and the total LC time was 141 min. The LTQ-Orbitrap XL was set to scan the precursors in the Orbitrap at a resolution of 60,000, followed by data-dependent MS/MS of the top 4 precursors.

The raw LC-MS/MS data was then submitted to Sorcerer Enterprise (Sage-N Research Inc.) for protein identification against the IPI rat protein database, which contains semi-trypsin peptide sequences with the allowance of up to 2 missed cleavages and precursor mass tolerance of 50.0 p.p.m. A molecular mass of 57 Da was added to all cysteines to account for carboxyamidomethylation. Differential search includes 16 Da for methionine oxidation. The search results were viewed, sorted, filtered, and statically analysed using PeptideProphet and ProteinProphet (ISB). The minimum trans-proteomic pipeline (TPP) probability score for proteins and peptides was set to 0.95, respectively, to assure a TPP error rate of lower than 0.01. The example MS/MS spectrum R.GLCVDALIELSDENADWK.L was identified as Fstl1 (Fig. 2a). Peptide probability = 1.0, Xcorr = 6.276, delta Cn = 0.471.

Automated *in vitro* cell proliferation and cell death assay. Cells (mCME^{ESC} and NRVC) were incubated with EdU (details of dosage and length of exposure are specified in figure legends) in a 384-well plate format, and were fixed for 2 h in 4% PFA, washed in PBS and stained for EdU using Click-it EdU assay kit (Life Technologies). The cells were then washed in PBS, immunostained with an α -actinin antibody (Sigma, A7811, 1:500) to identify cardiomyocytes and stained with DAPI (4',6-diamidino-2-phenylindole, 1:10,000) to identify nuclei. The plates were then imaged using InCell 1000 system (GE Healthcare) and automatically analysed in Developer Toolbox (GE Healthcare) as described³¹. Ratios of EdU⁺/ α -actinin⁺ nuclei and α -actinin⁺ nuclei were generated for the percentage of cardiomyocyte incorporated EdU in the chromosomal DNA.

Similarly, cells (mCME^{ESC} and NRVC) in 384-well plate format were fixed for 2 h in 4% PFA, washed in PBS, and were immunostained with pH3 antibody (Millipore 06-570, 1:200) for nuclei in mitosis, or aurora B (Millipore 04-1036, 1:200) for cytokinesis, or TUNEL (Roche) for cell death, and α -actinin antibody (Sigma, A7811, 1:500) for cardiomyocytes and DAPI (1:10,000) for nuclei. The same imaging and analysis were done for pH3 staining as the EdU assays, and the aurora B⁺, α -actinin⁺ double positive cells were manually counted. The percentages of pH3⁺, α -actinin⁺ double positive nuclei, aurora B⁺, α -actinin⁺ double positive cells, and TUNEL⁺, α -actinin⁺ double positive nuclei relative to the total number of α -actinin⁺ cell nuclei were calculated to determine the percentages of cardiomyocytes undergoing mitosis, cytokinesis and apoptosis, respectively. Calcium Imaging. Contractile calcium transients were recorded using a Kinetic Image Cytometer (KIC, Vala Sciences) using Fluo4 NW calcium indicator (Life Science). Data was processed using Cytoseer software containing the KIC analysis package (Vala Sciences) as described³².

Compressed collagen gel for use as an engineered epicardial patch. Highly hydrated collagen gels, used as cardiac patch in this study, were produced by adding 1.1 ml 1 \times DMEM (Sigma, MO, US) to 0.9 ml of sterile rat tail type I collagen solution in acetic acid (3.84 mg ml⁻¹, Millipore, MA, US). The resulting 2 ml collagen-DMEM mixture was mixed well and neutralized with 0.1 M NaOH (~50 μ l). The entire process was conducted on ice to avoid premature gelation of collagen. In the case of patches containing epicardial factors, the EMC culture media was collected as above and 0.6 ml of that was mixed with 0.5 ml DMEM. The collagen solution (0.9 ml) was then distributed into the wells of 24-well plates (15.6 mm in diameter) and placed in a tissue culture incubator for 30 min at 37 °C for polymerization. For pig studies, 6.8 ml of collagen was mixed with 8.2 ml DMEM to obtain a 15 ml solution that was then cast into a 6-cm Petri dish (area = 28.3 cm²). Plastic compression was performed as described previously^{33,34}. Briefly, as cast, highly hydrated collagen gels (at ~0.9 and 15 ml volumes for the mice and swine study, respectively) underwent unconfined compression via application of a static compressive stress of ~1,400 Pa for 5 min (see refs 33, 35 for details), resulting in ~98–99% volume reduction (Fig. 1n). The elastic modulus of the compressed collagen, aimed to approximate that of the embryonic epicardium which is optimal for contractility of immature cardiomyocytes (see ref. 36), was assessed by atomic force microscopy (AFM) in nano-indentation mode, using a force trigger that resulted in a minimal local strain of less than 10% (indentation of ~100 nm) to minimize the effect of substrate-related artefacts. A custom-made flat AFM tip was manufactured using focused ion beam milling and used to probe the stiffness of the gels by scanning areas of 90 μ m \times 90 μ m. Histogram of the distribution of measured micro stiffness of the patch is compared with the range of elasticity reported for common scaffolding biomaterials³⁷, and previously described³⁶ optimal range of elasticity to maximize myocyte contractility (Fig. 1o, $n = 3$).

Myocardial infarction and application of the epicardial patch. Permanent LAD occlusion (MI) Male 10–12-week-old C57BL/6J mice were purchased from Jackson Laboratories (Bar Harbour, ME, USA). Fstl1-TG mice used in MI

experiments are C57BL/6 background, female and male mice aged 12–15 weeks old. Mice were anaesthetized using an isoflurane inhalational chamber, endotracheally intubated using a 22-gauge angiocatheter (Becton, Dickinson Inc., Sandy, Utah) and connected to a small animal volume-control ventilator (Harvard Apparatus, Holliston, MA). A left thoracotomy was performed via the fourth intercostal space and the lungs retracted to expose the heart. After opening the pericardium, a 7-0 suture was placed to occlude the left anterior descending artery (LAD) ~ 2 mm below the edge of the left atrium. Ligation was considered successful when the LV wall turned pale (Fig. 1p). In the case of experimental groups treated with patch, immediately after the ligation, prepared collagen patch was sutured (at two points) onto the surface of ischaemic myocardium (Fig. 1g). The patch size used was ~ one-third of the 15.6 mm-diameter collagen gel. Animals were kept on a heating pad until they recovered. Another group of mice underwent sham ligation; they had a similar surgical procedure without LAD ligation. A minimum number of $n = 8$ was used in each study group.

Ischaemia reperfusion (I/R) Male C57BL/6, aged 10 to 11 weeks, were anaesthetized and intubated as described above. A left lateral thoracotomy was then performed. Pericardium was gently pulled off and an 8-0 Nylon suture (Ethicon, Inc. Johnson & Johnson Co., USA) was used to ligate the left anterior descending coronary artery against a PE10 tubing, which was removed after 30 min occlusion. Successful performance of coronary artery occlusion was verified by visual inspection (by noting the development of a pale colour in the distal myocardium upon ligation). The chest was then closed using 7-0 sutures around adjacent ribs, and the skin was closed with 6-0 suture. Buprenorphine was administered subcutaneously for a minimum of 1 day at BID dosing. For the animal group treated with patch, a second thoracotomy was performed one week post the incidence of I/R and the prepared collagen patch was sutured (at two points) onto the surface of ischaemic myocardium. Sham-operated controls consisted of age-matched mice that underwent identical surgical procedures (two thoracotomies) with the exception of LAD ligation (Extended Data Fig. 4).

Echocardiography. *In vivo* heart function was evaluated by echocardiography at 2 weeks (Figs 1r, s and 3b, c) and Extended Data Fig. 2 h–j), 4 weeks (Fig. 3b, c, e and Extended Data Fig. 2h–j), and 2 and 3 months (Fig. 3b) after LAD ligation. Two-dimensional (2D) analysis was performed on mice using a GE Vivid 7 ultrasound platform (GE Health Care, Milwaukee, WI) equipped with 13 MHz transducer. The mice were sedated with isoflurane (100 mg per kg, inhalation), and the chest was shaved. The mice were placed on a heated platform in the supine or left lateral decubitus position to facilitate echocardiography. 2D clips and M-mode images were recorded in a short axis view from the mid-left ventricle at the tips of the papillary muscles. LV internal diameter (LVID) and posterior wall thickness (LVPW) were measured both at end diastolic and systolic. Fractional shortening (FS, %) and ejection fraction (EF, %, via extrapolation of 2D data) were calculated from LV dimensions in the 2D short axis view. A minimum number (n) of 8 mice per experimental group was used for the echo evaluations. Measurements were performed by two independent groups in a blind manner. In ischaemia reperfusion study, *in vivo* heart function was evaluated pre-surgery (baseline), 1 week after the incidence of I/R, and two and four weeks post-implantation (Extended Data Fig. 4a–d).

***In vivo* delayed-enhanced magnetic resonance imaging (DEMRI).** To prepare for scanning, induction of anaesthesia was accomplished with 2% and maintained with 1.25–1.5% isoflurane with monitoring of the respiratory rate. ECG leads were inserted subcutaneously to monitor the heart rate while the body temperature was maintained at 37 °C. Using 3T GE Signa Excite clinical scanner with a dedicated mouse coil (RAPID MR International, Germany), functional parameters were recorded on weeks 1 and 4 after treatment. The following sequences were performed for MRI acquisitions: (1) DEMRI was performed following i.p. injection of 0.2 mmol per kg gadopentetate dimeglumine (Magnevist, Berlex Laboratories) using gated fGRE-IR sequences with FOV 3.4 cm, slice thickness 0.9 mm, matrix 128 × 128, TE 5 ms, TI 150–240 ms, and FA 60°; and (2) cardiac MRI of volumes were performed using fSPGR with FOV 7 cm, slice thickness 0.9 mm, matrix 256 × 256, TE 5.5 ms, and FA 30. Coronal and axial scout images were used to position a 2-dimensional imaging plane along the short axis of the left ventricular (LV) cavity (Extended Data Fig. 3h–j).

Histology, immunohistochemistry and immunofluorescent staining. Histological analysis (Mason's trichrome staining) was performed according to standard protocols for paraffin embedded samples. For immunohistochemistry and immunofluorescent staining, embedded hearts were sectioned at a thickness of 7 µm, unless described otherwise. Antibodies used were as follows: 1:200 α -actinin (Sigma, A7811), 1:300 α -smooth muscle actin (Sigma A2547) 1:100 phospho-Histone3 (rabbit Millipore 06-570), 1:300 phospho-Histone3 (mouse Abcam ab14955) 1:100 WT1 (Abcam, ab15249), 1:100 (Fig. 1l–m) and 1:250 (Figs 3m, n and 5n) aurora B (Millipore 04-1036), 1:100 Tnni2 (DSHB, Ct3), 1:100 Tnni3 (Abcam, ab56357), 1:200 PCMI (Sigma-Aldrich HPA023370), 1:200

FSTL1 (R&D MAB17381). At least 5 sections per heart were used per staining for Mason's Trichrome staining and 3 sections per heart per staining for immunohistochemistry and immunofluorescent staining, respectively. HRP Anti-rat secondary antibody (Jackson ImmunoResearch 712-036-153, 1:500) was used for immunohistochemistry, and respective fluorescent secondary antibodies (Life Technologies 1:200) were used for immunofluorescent staining. The Trichrome staining and immunohistochemistry images were taken using an upright Zeiss microscope and dissection scopes. The fluorescent images were taken using Apotome Optical Sectioning (Zeiss). An inclusion criterion for the patch engraftment was that the patch covered > 70% of the infarct (controlled by histology). TUNEL assay (Roche 11684795910) and EdU assay (Life Technologies C10337) were performed as instructed.

Lineage tracing experiments. Epicardial lineage labelling was achieved by oral delivery of tamoxifen (4 mg) in eight-week old $Wt1^{creERT2/+}; Rosa26^{RFP/+}$ mice with C57BL/6J and ICR mixed background (delivered 6 times for duration of 3 weeks and stopped 1 week before MI). Hearts were collected at 2 weeks after MI. Immunostaining of RFP for $Wt1$ lineage cells, Fstl1 and Tnni3 shows that Fstl1 is absent in epicardial cells and their derivatives, but abundant in the myocardium after MI (Fig. 2l).

Cardiomyocyte lineage labelling was achieved by injecting 4-OH tamoxifen intraperitoneally into eight-week old $Myh6^{mERcreER}; Rosa26^{Z/EG}$ mice¹⁷ of C57BL/6 background at a dose of 20 mg per kg per day for 2 weeks, and stopped 1 week before harvesting cardiomyocytes (Extended Data Fig. 7a–f), or MI operation and patch grafting. 4 weeks after MI, the animals were collected for immunostaining (Fig. 3q–v).

TTC staining. At day 2 post MI/patch treatment, the mouse hearts from all four groups were harvested and sectioned perpendicularly to the long axis into four sections (approximately 2 mm thick). The sections were placed in the wells of a 12-well cell culture plate and incubated with 1% 2,3,5-triphenyltetrazolium chloride (TTC, Sigma-Aldrich) solution for 15 min at 37 °C. Subsequently section were washed with PBS and visualized using a stereomicroscope and photographed with a digital camera (Extended Data Fig. 6a, b).

Vessel counting. Blood vessel density parameters were measured from histological sections of heart samples stained for von Willebrand factor (vWF) as a marker of endothelial cells in the vessel wall. Up to 60 sections were analysed for each treatment group (4 mice in each group). Analysis was performed using ImageJ to calculate: (1) the total luminal area of blood vessels, and (2) the number of vessels that stained + for the vWF. In each case, a histogram of the vessel parameters as a fraction of total surface area analysed was obtained and the mid-values plotted for each treatment group. Statistical significance ($P < 0.05$) of the differences from sham group was determined by one-tailed ANOVA (Fig. 3f–i).

Cardiomyocyte proliferation quantification *in vivo*. Data collected from 5–7 hearts in each group (7 for MI plus patch with FSTL1, 5 for Sham, MI-only and MI plus patch) with 3 different cross-sections (each section covered the infarct, patch, and separated by 250 µm, between 1–2 mm from the apex) counted exhaustively for total pH3⁺/ α -actinin⁺, aurora B⁺/ α -actinin⁺, and pH3⁺/PCMI⁺ cells in each section, and normalized to myocardium area quantified by trichrome staining of immediate adjacent section (Fig. 5j–p, and Extended Data Fig. 5).

Enzyme-linked immunosorbent assay. In order to assess the FSTL1 retention within the engineered patch system *in vitro*, collagen scaffolds laden with FSTL1 (5 µg ml^{−1}) were immersed in PBS and shaken for various times (0, 12 h, 1 day, and 21 days) at 4 °C and the FSTL1 concentration was determined using Enzyme-linked Immunosorbent Assay kit (USCN Life Science, Inc., Houston, USA). The detection limit for this technique was 0.50 ng ml^{−1}. Scaffolds were pretreated with 1 mg ml^{−1} collagenase type I (Sigma Aldrich, MO, US) and 5 mg ml^{−1} hyaluronidase (Sigma Aldrich, MO, US) dissolved in phosphate buffered saline for 5 min followed by centrifugation at 5,000 g for 20 min.

Aliquots of 100 µl of the collected samples were added to the 96-well plates and incubated for 2 h at 37 °C. Then, 100 µl of the prepared detection reagent A were added to the wells followed by 1 h incubation at same temperature. After aspiration and washing 3 times, 100 µl of the prepared detection reagent B was added to the wells and incubated for 30 min at 37 °C. After aspiration and washing 5 times, 90 µl of substrate solution was added to the wells following by incubation for 25 min at 37 °C. 50 µL of stopping solution was added to the wells and the absorbance of each well was read at 450 nm, immediately. The concentration of FSTL1 was defined using standard curve of the standard solutions. The test was performed 4 times (Extended Data Fig. 3a).

Application of the patch in a swine model of ischaemia-reperfusion. The swine study was performed by inflation of a percutaneous coronary angioplasty dilation catheter to occlude the LAD in Yorkshire pigs (45 days old). Occlusion time of 90 min was followed by fully reperfusion to mimic the clinical MI disease model.

One week after MI, a left thoracotomy was performed and the patch (6-cm diameter) was sutured onto the infarct. Animal groups included: sham controls, I/R with no treatment ($n = 3$), I/R treated with patch alone (I/R plus patch, $n = 1$), and I/R treated with patch laden with FSTL1 (I/R plus patch with FSTL1, $n = 2$). EdU delivery: 250 mg per week EdU was infused into circulation during the 4-week time course of study (week 1 to week 5 post I/R), using osmotic mini pumps (Fig. 5).

Animal compliance. The procedures involving animal use and surgeries were approved by the Stanford Institutional Animal Care and Use Committee (IACUC). Animal care and interventions were provided in accordance with the Laboratory Animal Welfare Act (C57BL/6J) wildtype mice (Figs 1p–t, 2h–k, 3a–p, 3–6) *Myh6^{mERcre};Rosa26^{Z/EG}*. C57BL/6J mice (Fig. 3q–v, Extended Data Fig. 7a–f) Yorkshire pigs (Fig. 5)).

The study protocol was approved by the Institutional Animal Care and Use Committee (IACUC) of Boston University (wild-type and Fstl1-TG C57BL/6J mice, (Fig. 3c and Extended Data Fig. 2)).

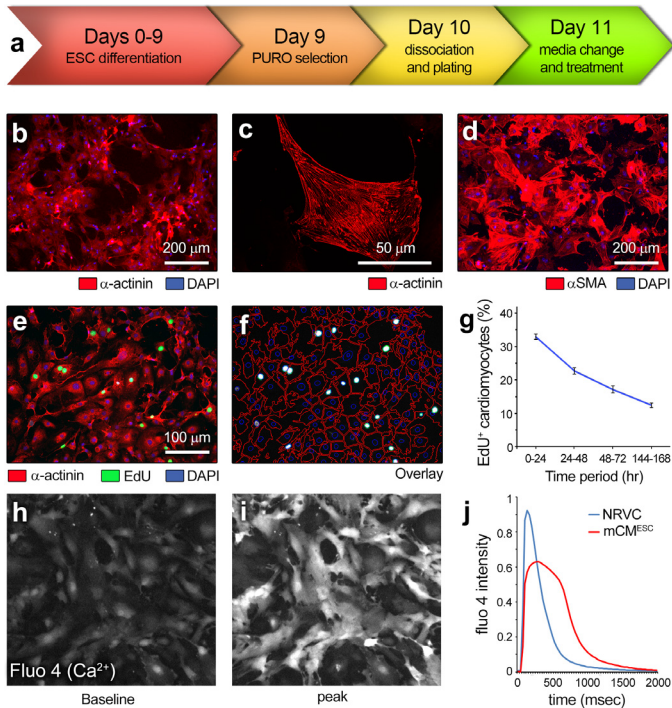
Mice were used in accordance with the guidelines of the Institutional Animal Care and Use Committee (IACUC) of the Institute for Nutritional Sciences, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences (*Tnt-cre;Rosa26^{mTmG/+}; Wt1^{creERT2/+};Rosa26^{mTmG/+}* mice (C57BL/6J and ICR mixed background) (Fig. 1j–m) *Wt1^{creERT2/+};Rosa26^{RFP/+}* mice (C57BL/6J and ICR mixed background) (Fig. 2l)

All animal study was approved by the Institutional Animal Care and Use Committee (IACUC) of Sanford-Burnham-Prebys Medical Discovery Institute. All animal procedures performed conform the NIH guidelines (Hsd:s.d. rats (Fig. 4g, m, n and Extended Data Fig. 7g–j)).

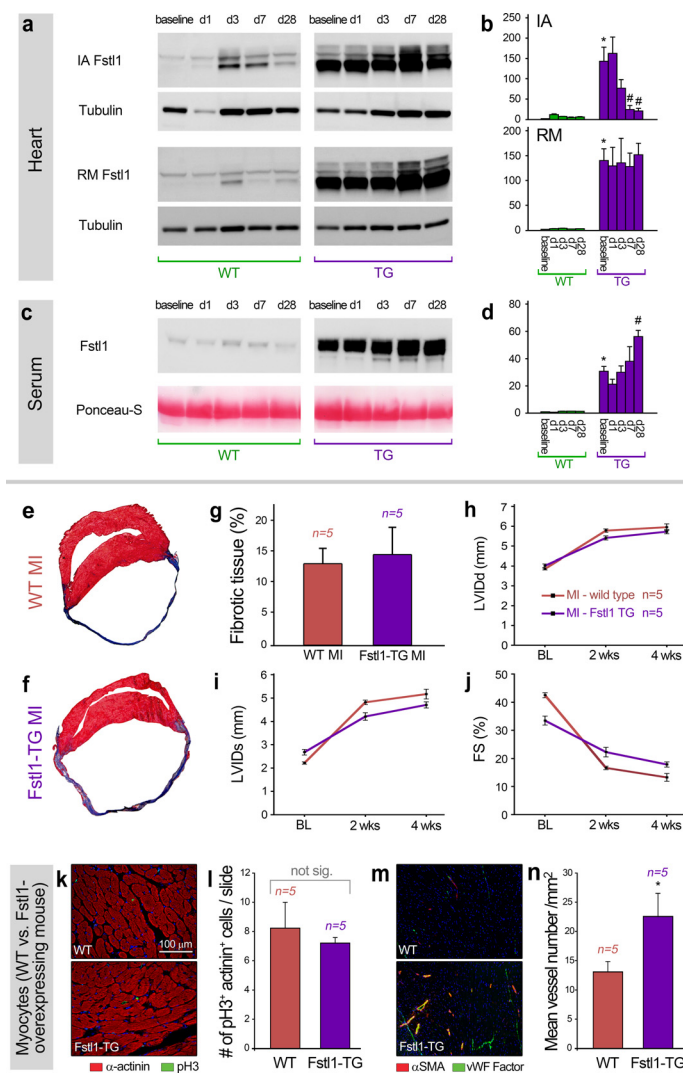
Statistical analysis. The number of samples (n) used in each experiment is recorded in the text and shown in figures. All *in vitro* experiments have been done at least twice independently. Gene expression experiments have been done 3 times independently and EdU proliferation assays and cell size measurement have been done more than 10 times independently. Sample size was not pre-determined, with retrospective analysis of significantly different results in most *in vitro* studies using Gpower 3.1 produces power > 0.8 . Sample sizes for animal studies were estimated. Animals which did not survive up to 4 weeks after surgery were excluded from functional and histological studies. Randomization was not applied. Blinding to group allocation was practiced between animals surgery and results analysis of mouse myocardial infarction experiments. The values presented are expressed as means \pm s.e.m. The rationale to use means \pm s.e.m. instead of s.d. is that s.e.m. quantifies uncertainty in an estimate of the mean whereas s.d. indicates dispersion of the data from mean. In other words, the SEM

provides an estimate of the reported mean value, while the s.d. gives an idea of the variability of single observations. Normal distribution were tested and confirmed in automatic analysis of mCMs^{ESC} (Figs 1c, g, i, 2d, f, g and 4d, e, j,–n and Extended Data Figs 1 and 7h, j). We did not estimate variations in the data. The variances are similar between the groups that are being statistically compared. One-way ANOVA with multiple comparisons (Fig. 1r, 3 and Extended Data Figs 2b, d and 4, 5, 6) and Student's *t*-test (Fig. 1a–m, 2 and 4 and Extended Data Fig. 2e–n and 7) were used to test for statistical significance ($P < 0.05$). Survival curve were generated by Kaplan–Meier method using PRISM (GraphPad) and Log-rank (Mantel–Cox) test was used to test the significant differences between the survival of mice in different conditions (Fig. 3a).

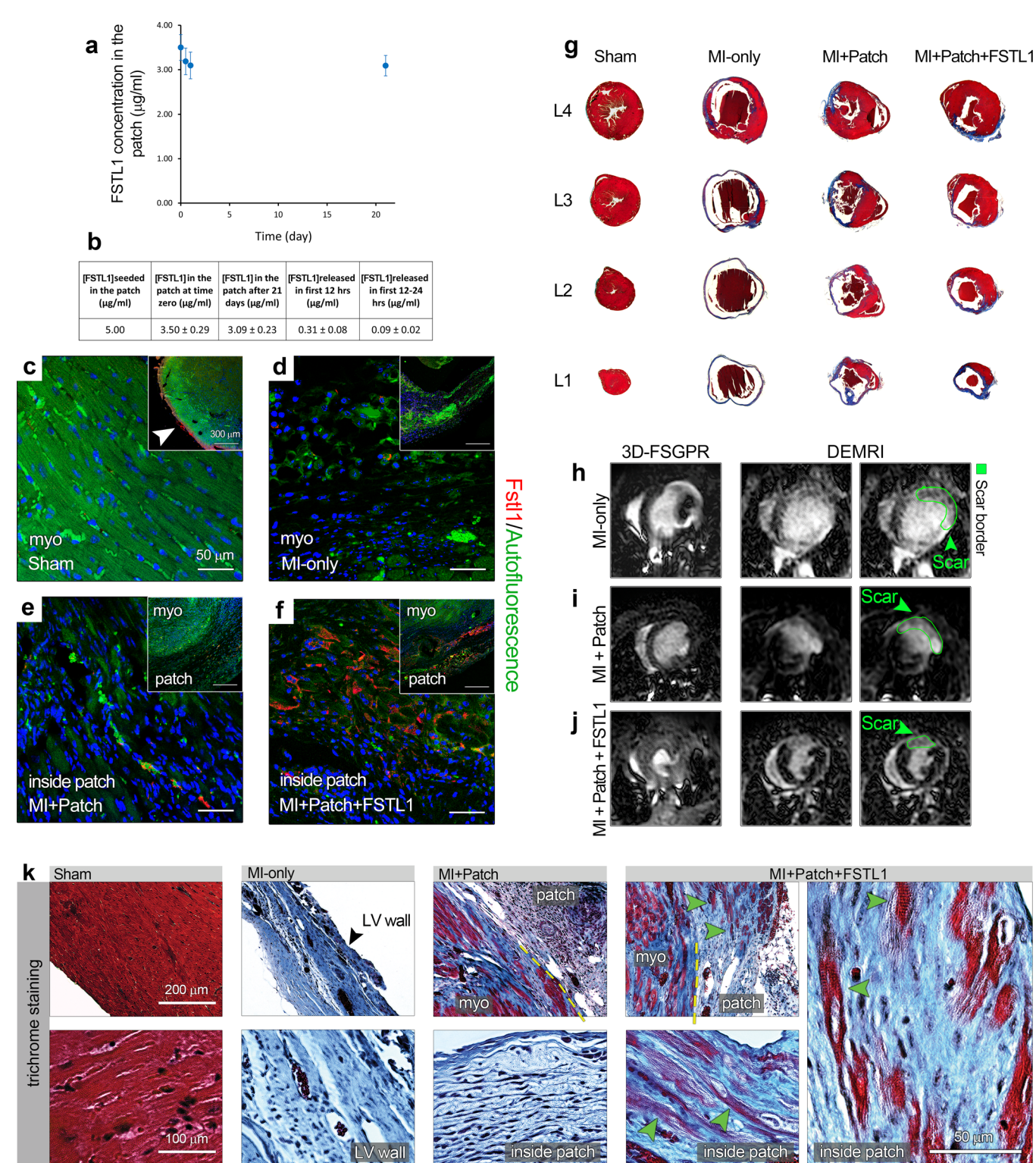
27. Eid, H. *et al.* Role of epicardial mesothelial cells in the modification of phenotype and function of adult rat ventricular myocytes in primary coculture. *Circ. Res.* **71**, 40–50 (1992).
28. Kita-Matsuo, H. *et al.* Lentiviral vectors and protocols for creation of stable hESC lines for fluorescent tracking and drug resistance selection of cardiomyocytes. *PLoS ONE* **4**, e5046 (2009).
29. Fajardo, G. *et al.* Deletion of the β 2-adrenergic receptor prevents the development of cardiomyopathy in mice. *J. Mol. Cell Cardiol.* **63**, 155–164 (2013).
30. Sakaue-Sawano, A. *et al.* Visualizing spatiotemporal dynamics of multicellular cell-cycle progression. *Cell* **132**, 487–498 (2008).
31. Bushway, P. J. & Mercola, M. High-throughput screening for modulators of stem cell differentiation. *Methods Enzymol.* **414**, 300–316 (2006).
32. Cerignoli, F. *et al.* High throughput drug risk assessment in human cardiomyocytes by kinetic image cytometry. *J. Pharm. Toxicol. Methods* **66**, 246–256 (2012).
33. Serpooshan, V. *et al.* Reduced hydraulic permeability of three-dimensional collagen scaffolds attenuates gel contraction and promotes the growth and differentiation of mesenchymal stem cells. *Acta Biomater.* **6**, 3978–3987 (2010).
34. Serpooshan, V., Muja, N., Marelli, B. & Nazhat, S. N. Fibroblast contractility and growth in plastic compressed collagen gel scaffolds with microstructures correlated with hydraulic permeability. *J. Biomed. Mater. Res. A* **96**, 609–620 (2011).
35. Abou Neel, E. A., Cheema, U., Knowles, J. C., Brown, R. A. & Nazhat, S. N. Use of multiple unconfined compression for control of collagen gel scaffold density and mechanical properties. *Soft Matter* **2**, 986–992 (2006).
36. Engler, A. J. *et al.* Embryonic cardiomyocytes beat best on a matrix with heart-like elasticity: scar-like rigidity inhibits beating. *J. Cell Sci.* **121**, 3794–3802 (2008).
37. Venugopal, J. R. *et al.* Biomaterial strategies for alleviation of myocardial infarction. *J. Royal Soc. Interface* **9**, 1–19 (2012).
38. Clement, S. *et al.* Expression and function of alpha-smooth muscle actin during embryonic-stem-cell-derived cardiomyocyte differentiation. *J. Cell Sci.* **120**, 229–238 (2007).



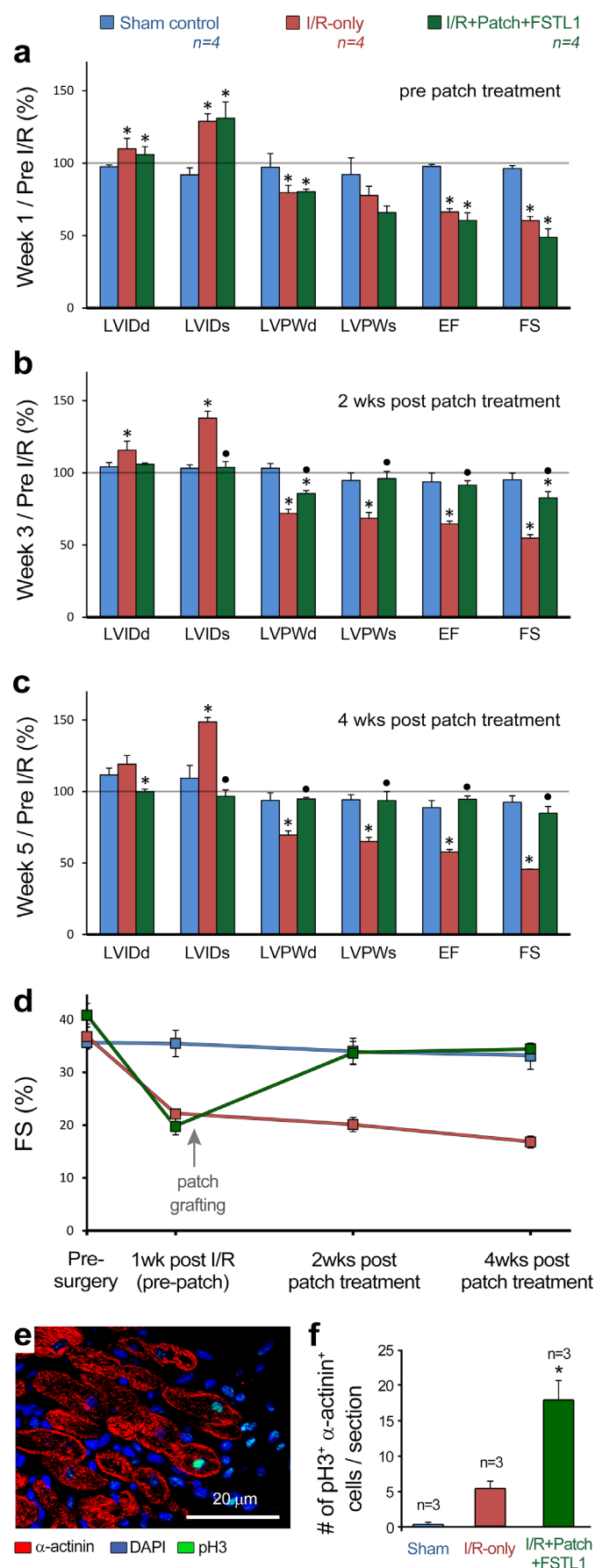
Extended Data Figure 1 | Characterization of mCMs^{ESC} cells used in this study. **a**, Schematic time-line of cell preparation and treatment. **b–d**, Immunostaining of α -actinin of mCMs^{ESC}, showing that the majority of the cells are α -actinin⁺ (**b**), and the α -actinin lacks striation structures (**c**). **d**, Immunostaining of α -smooth muscle actin (α SMA) of mCMs^{ESC}, showing the majority of the cells are α SMA⁺, unlike mature cardiomyocytes with no SMA expression³⁸. **e, f**, Automatic detection of EdU incorporation in mCMs^{ESC}. Captured image of mCMs^{ESC} treated with 10 μ g ml⁻¹ EdU for 24 h, stained with EdU, α -actinin and DAPI using InCell 1000 (General Electric) (**e**). Overlay of masks of EdU, α -actinin and DAPI channels with automatic detection software (**f**). **g**, EdU incorporation profile of mCMs^{ESC} over time. mCMs^{ESC} are treated with 10 μ g ml⁻¹ EdU for 24 h at time 0 h, 24 h, 48 h, and 144 h. The percentage of EdU⁺/ α -actinin⁺ cardiomyocytes of all α -actinin⁺ cardiomyocytes is calculated for each time period. Note the decrease of EdU incorporation rate over time. **h, i**, Fluo 4 calcium images of mCMs^{ESC}, with baseline background image (**h**) and peak image (**i**). **j**, Comparison of representative calcium transients of mCMs^{ESC} (red) and neonatal rat ventricular cardiomyocytes (NRVC, blue). Note the reduced amplitude, slower rate of up and down strokes, and elongated duration of the calcium transient in mCMs^{ESC} compared to NRVC, suggesting immature calcium handling in mCMs^{ESC}. In all experiments, FSTL1 was added one day after plating of the mCMs^{ESC} (time 0–24 in this figure).



Extended Data Figure 2 | Myocardial overexpression of Fstl1 (Fstl1-TG) mice after permanent LAD ligation. **a–d**, Fstl1 protein expression kinetics after myocardial infarction. Fstl1-TG mice (C57/BL6 background) and littermate wild-type (WT) mice underwent LAD ligation. Heart tissue and serum were collected at baseline, day 1, day 3, day 7 and day 28 after surgery. Fstl1 protein levels in ischaemic area (IA) and remote area (RM) of heart were analysed by western blotting (**a**). Fstl1 expression expressed relative to tubulin levels is reported (**b**). Fstl1 serum levels were analysed by western blotting (**c**). Also shown in Ponceau-S staining to indicate equal loading of serum. Quantification of serum Fstl1 level is shown in (**d**). $n > 3$ in all groups. * $P < 0.05$ compared to WT baseline, # $P < 0.05$ compared to Fstl1-TG baseline. ANOVA was used for statistical significance ($P < 0.05$). **e–j**, Morphometric and functional response of Fstl1-TG mice to permanent LAD ligation at long-term. Representative Masson's trichrome staining of WT (**e**) and Fstl1-TG (**f**) 4 weeks after MI. Quantification of content in fibrotic tissue at week 4 after MI (**g**). Echocardiographic measurement of left ventricular internal dimension in systole (LVIDs) (**h**), and left ventricular internal diameter in diastole (LVIDd) (**i**) at weeks 2 and 4 after MI. Echocardiographic determination of fractional shortening (FS%) in the indicated genotypes at 2 and 4 weeks after MI (**j**). **k–n**, Double immunofluorescent staining of α -actinin (cardiomyocytes) and pH3 (mitosis) (**k**) and α -actinin (cardiomyocytes) and von Willebrand factor (vascular endothelial cells) (**m**) in the Fstl1-TG and WT mice, quantified in (**i**, **n**). $n = 5$, * $P < 0.05$ indicates significantly different from WT.

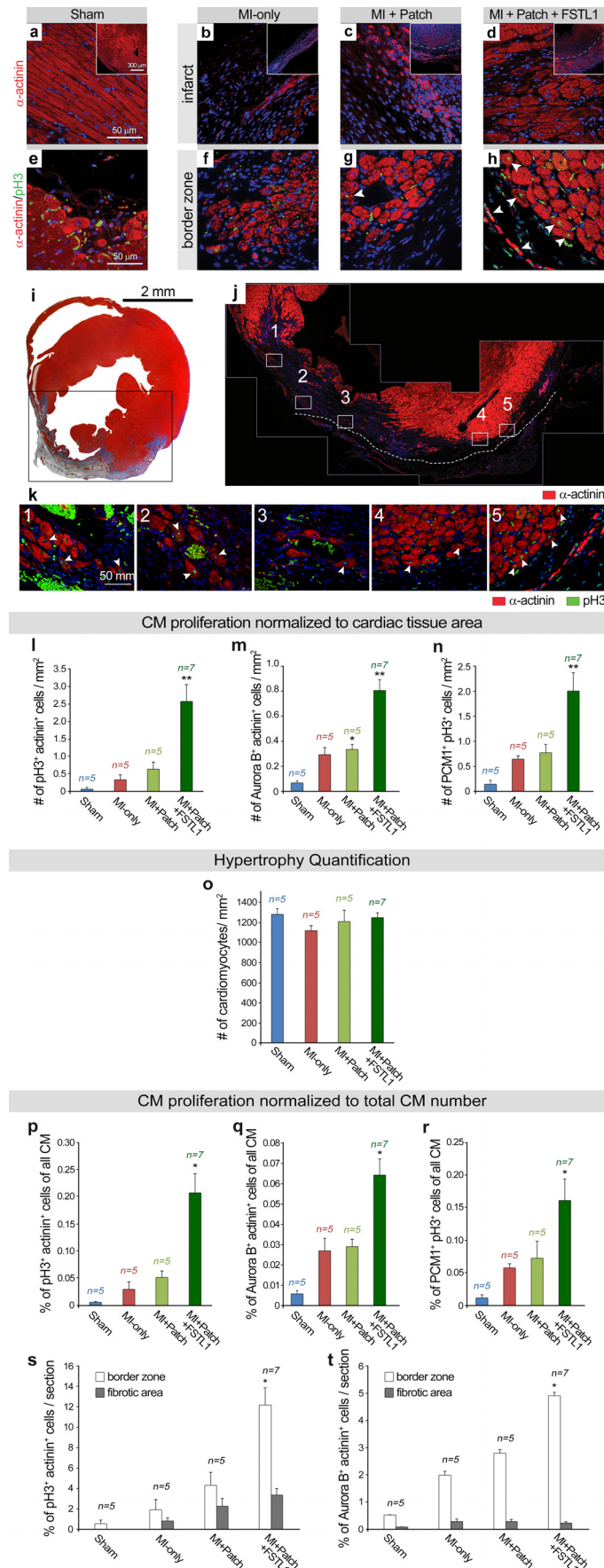


conditions (sham, MI only, MI with patch and MI plus patch with FSTL1) 4 weeks after MI. Note the severe fibrosis in MI only condition, and reduced fibrosis in MI plus patch condition, and further reduction in MI plus patch with FSTL1 condition, quantified in Fig. 3d. **h–j**, Representative MRI images from the mouse MI only, MI plus patch and MI plus patch with FSTL1 treatment groups showing the 3D-FSPGR (fast spoiled gradient-echo) images and the delayed enhancement images using gadolinium contrasting agents, confirming a reduction in infarct area (demarcated in green) and preserved contractility (Supplementary Videos 3–5). **k**, Trichrome staining of infarct and border zone of the indicated treatments demonstrates the integration of the patch with the host tissue and massive patch cellularization by the native cardiac cells. Observe the abundant muscle (red) inside the patch and in the border zone of the patch with FSTL1 treated animals (three right panels, green arrowheads).



Extended Data Figure 4 | Analysis of patch with FSTL1 function in the mouse model of ischaemia/reperfusion (I/R) with delayed patch grafting.
a–c, Heart function evaluation for sham, I/R, and I/R treated with patch with FSTL1, at end-diastolic and systolic, pre-grafting (**a**, 1 week post-injury), 2 weeks post patch implantation (**b**), and 4 weeks post grafting (**c**). Values were normalized by dividing to pre-surgery baseline values for each individual animal. **d,** Absolute values of fractional shortening (FS, %) at different times pre and post I/R as evaluated by echocardiography of mice from

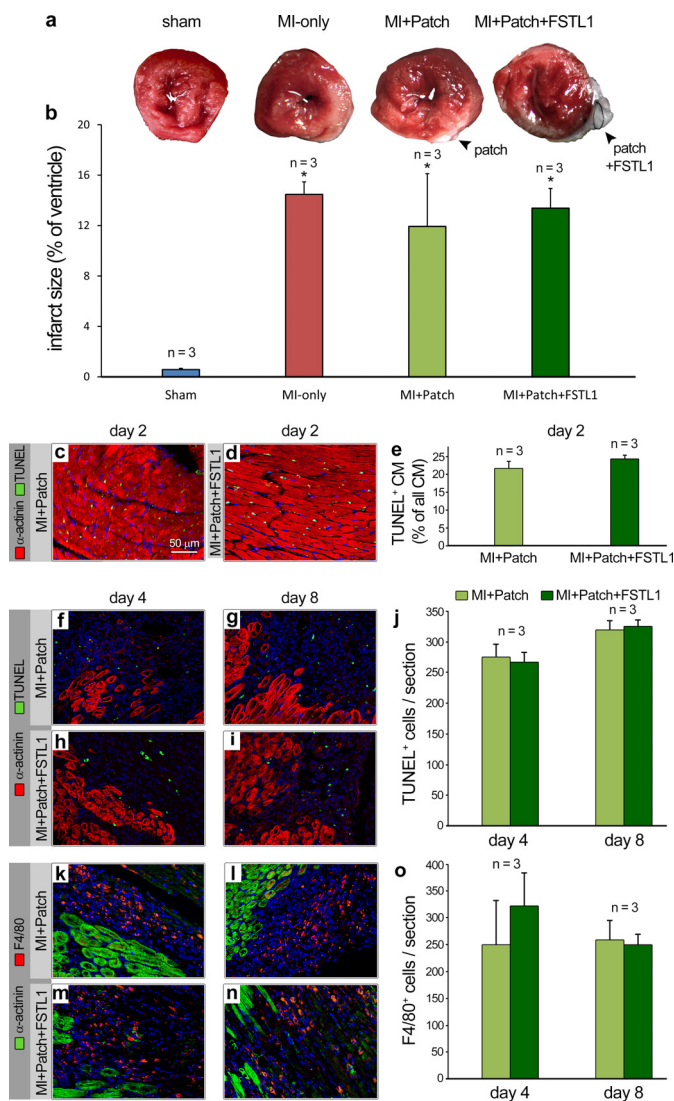
a–c. Abbreviations same as in Fig. 3. * $P < 0.05$ compared to sham and black circle $P < 0.05$ compared to I/R. **e,** Co-immunofluorescence staining of DNA duplication marker phospho-Histone3 Ser10 (pH3, green) and α -actinin (red) in the border zone of patch with FSTL1 treated heart 4 weeks after MI. **f,** Quantification of incidence of pH3⁺, α -actinin⁺ double positive cells in the 3 experimental groups. Data collected from 3 hearts in each group with 3 different cross sections counted for total pH3⁺, α -actinin⁺ cells in each heart. * $P < 0.05$ indicates statistically different from all other groups.



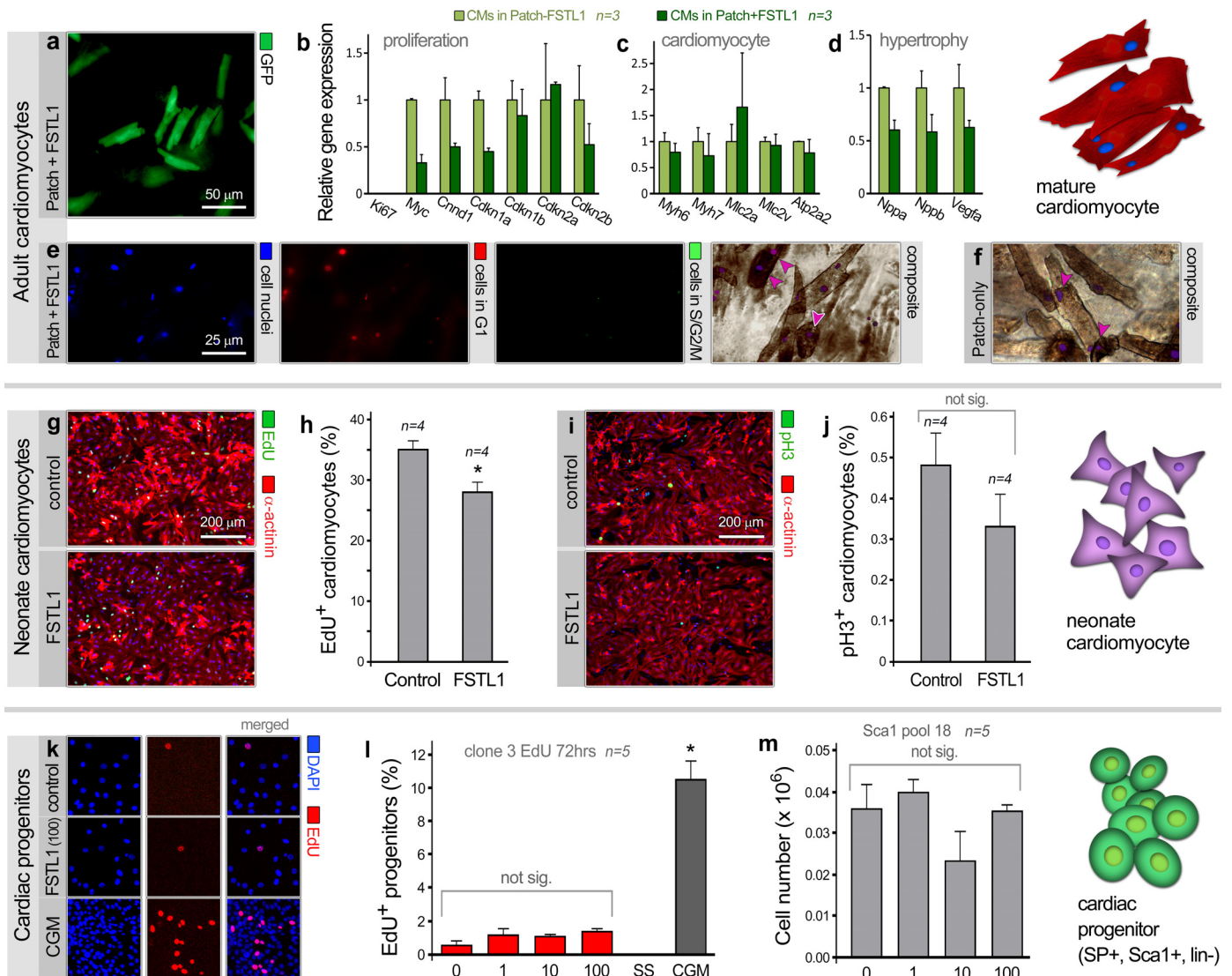
Extended Data Figure 5 | Representative images and quantification of cardiomyocyte proliferation *in vivo* after patch with FSTL1 treatment.

a–h, Immunostaining of the cardiomyocyte marker α -actinin (red) in the infarct area (**b–d**) and co-immunofluorescence staining of DNA duplication marker phospho-Histone3 Ser10 (pH3, green) and α -actinin (red) in the border zone (**f–h**), in the 4 treatment groups analysed 4 weeks post-MI, compared to sham-operated animals (**a**, **e**). Insets in (**a–d**) show lower magnification images with broken lines demarcating the border between the patch and host tissues. Arrowheads in **g**, **h**, indicate α -actinin⁺ cardiomyocytes with pH3⁺ nuclei. **i–k**, Representative images of pH3⁺ cardiomyocytes in a patch with FSTL1 treated heart. Masson's Trichrome staining of a heart after MI 4 weeks treated with patch with FSTL1 (**i**). The adjacent slide was stained for α -actinin in **j**, corresponding to the black box area with infarction and the patch in **i**. The spotted line in **j** indicates the boundary between the heart and the patch. The adjacent slide was stained for α -actinin and pH3, and all α -actinin⁺, pH3⁺ double positive cardiomyocytes found were shown in **k** (white arrowhead), with each image corresponding to the area in numbered white boxes in **j**. **l–n**, Quantification of cardiomyocyte proliferation measured in 3 cross sections covering the infarct, patch, and separated by 250 μ m, between 1–2 mm from the apex in each heart (Fig. 3j). Data collected from 5–7 hearts in each group with the 3 cross-sections counted exhaustively for incidence of

α -actinin⁺ cells positive for pH3 (**l**), midbody-localized aurora B kinase between α -actinin⁺ cells (**m**), and double-positive cells for pH3 and the nuclear cardiomyocyte maker PCMI (**n**), and normalized to myocardium area quantified by trichrome staining of immediate adjacent section. * $P < 0.05$ statistically different from sham. ** $P < 0.05$, statistically different from all other groups. **o**, Quantification of hypertrophy in all experimental groups, measured by counting cardiomyocytes in areas of intraventricular wall with perpendicular cross-sections of cardiomyocytes in all hearts analysed for cardiomyocyte proliferation. No significance were found between samples. **p–r**, Quantification of incidence of α -actinin⁺ cells positive for pH3 (**p**), midbody-localized aurora B kinase between α -actinin⁺ cells (**q**), and double-positive cells for pH3 and the nuclear cardiomyocyte maker PCMI (**r**) measured in **l**, **m**, to total number of cardiomyocytes, calculated using hypertrophic analysis results in **o**. * $P < 0.05$, statistically different from sham, ** $P < 0.05$; statistically different from all other groups. **s**, **t**, Quantification of incidence of α -actinin⁺ cells positive for pH3 (**s**) and midbody-localized aurora B kinase between α -actinin⁺ cells (**t**), separated by their localization in the border zone or infarcted area. Note the majority of proliferation quantified by both methods are located in the border zone, * $P < 0.05$, statistically different from all other groups.



Extended Data Figure 6 | Effect of implantation of patch with FSTL1 on apoptosis and inflammation. **a**, Representative TTC staining of day 2 post MI/patch treatment of all four groups (sham, MI, MI plus patch, MI plus patch with FSTL1). **b**, Quantification of area at risk comparing all 4 groups. Data collected from 4 hearts in each group, with 4 cross-sections, approximately 2 mm thick each, encompassing each heart. * $P < 0.05$, statistically different from the sham. **c, d**, Representative image of TUNEL assays (TUNEL, green, α -actinin, red) comparing hearts 2 days after MI with patch alone and patch with FSTL1. **e**, Quantification of TUNEL⁺, α -actinin⁺ in infarcted area, as percentage of total number of cardiomyocyte. No difference is observed between MI plus patch and MI plus patch with FSTL1 conditions. Data collected from 3 hearts in each group with 3 different cross-sections (same as in Fig. 3j). Ten 0.09 mm² images were taken from infarcted area of each section and counted for TUNEL⁺, α -actinin⁺ and total α -actinin⁺ cells. **f–j**, TUNEL staining for cell death and α -actinin staining for cardiomyocytes were performed on hearts treated with patch-only and patch with FSTL1 at day 4 and day 8 after MI (**f–i**). Minimal TUNEL⁺, α -actinin⁺ cells are detected while there are significant amount of TUNEL⁺, α -actinin⁺ cells. Quantification of all TUNEL⁺ nuclei showed no significant differences between patch and patch with FSTL1 treated hearts at both time points (**j**). **k–o**, Immunostaining of F4/80 for macrophages and α -actinin for cardiomyocyte were performed on the same hearts as in panels **a–d** (**k–n**). Quantification of F4/80⁺ cells showed no significant differences between patch and patch with FSTL1 treated hearts at both time points (**o**).



Extended Data Figure 7 | FSTL1 does not induce proliferation in adult and neonatal cardiomyocytes, or cardiac progenitor cells. **a–f**, Adult cardiomyocytes derived from mouse primary isolation. **a**, Visualization of GFP⁺ cardiomyocytes isolated from *Myh6^{mERcre/ER}; Rosa26^{Z/EG}* mice treated with 4-OH-tamoxifen (OH-Tam) in 3D-collagen patches. **b–d** Gene expression changes in adult cardiomyocyte treated with FSTL1, including proliferation (**b**), cardiac-specific (**c**), and hypertrophy (**d**) markers. Note no changes in expression of cardiac specific genes, no increase in cell cycle markers (consistent with undetectable Ki67 immunostaining), and decreased hypertrophy markers ($n = 3$). Cardiomyocytes were embedded within 3D patch were treated with FSTL1 (10 ng ml⁻¹) for duration of 7 days with media change every 2 days. **e**, Fucci assay in 3D-cultured adult cardiomyocytes, conducted 1 week after the 3D culture. **e**, Treatment with FSTL1 was performed for 7 days with media change every 2 days. **f**, Adult cardiomyocytes 3D-cultured control in absence of FSTL1. Note no detectable sign of cardiomyocytes in S/G2/M phases (GFP⁺) in either condition. Purple arrows point to purple-colored nuclei resulting from co-localization of Hoechst (blue) and G1 phase

Fucci (red) labelling. **g–j**, Primary neonatal rat ventricular cardiomyocytes (NRVC). **g**, Freshly isolated NRVCs stimulated with FSTL1 for 48 h with 10 $\mu\text{g ml}^{-1}$ EdU, and stained for α -actinin (red) and EdU (green). Percentages of EdU⁺/ α -actinin⁺ cardiomyocytes of all α -actinin⁺ cardiomyocytes are quantified (**h**). **i**, j, NRVCs stimulated with FSTL1 for 48 h, and stained for α -actinin (red) and pH3 (green). Percentages of pH3⁺/ α -actinin⁺ cardiomyocytes of all α -actinin⁺ cardiomyocytes are quantified (**j**). No increase of proliferation is found upon FSTL1 treatment. ($n = 4$) * $P < 0.05$, statistically different from control. **k–m**, Sca1⁺ progenitor cells¹⁸ were starvation-synchronized for 48 h and stimulated with FSTL1 or control growth medium for 72 h in presence of EdU. Clone 3 was obtained by clonal growth from the Lin-Sca1+SP fraction. Sca1 pool was obtained from lin-Sca1⁺ without clonal growth. **k**, EdU and DAPI staining of Sca1⁺ cells after 72 h treatment. **l**, Percentage of EdU⁺ Sca1⁺ cells after 72 h treatment. FSTL1 concentration: 0, 1, 10, 100 ng ml⁻¹. Abbreviation: SS, serum starvation; CGM, control growth medium. **m**, Number of Sca1⁺ cells after 72 h FSTL1 treatment ($n = 5$). No significant change is found upon FSTL1 treatment.

Extended Data Table 1 | Raw echocardiography values (average \pm s.e.m.) obtained at days 0 (baseline), 14, and 28 post treatment in a mouse model of permanent LAD ligation

	LVIDd (mm)	LVIDs (mm)	LVPWd (mm)	LVPWs (mm)	EF (%)	FS (%)
Sham Baseline, n=10	3.89 \pm 0.07	2.44 \pm 0.07	1.10 \pm 0.06	1.44 \pm 0.04	75.40 \pm 1.33	37.76 \pm 1.48
MI-only Baseline, n=10	3.94 \pm 0.06	2.28 \pm 0.09	1.29 \pm 0.09	1.68 \pm 0.11	78.53 \pm 1.68	43.50 \pm 1.82
MI+Patch Baseline, n=10	4.08 \pm 0.10	2.50 \pm 0.16	1.07 \pm 0.03	1.49 \pm 0.12	77.75 \pm 2.67	41.54 \pm 1.98
MI+Patch+CM Baseline, n=10	3.84 \pm 0.09	2.28 \pm 0.07	1.19 \pm 0.04	1.60 \pm 0.05	77.63 \pm 0.71	40.72 \pm 0.66
MI+Patch+FSTL1 Baseline, n=10	4.04 \pm 0.14	2.42 \pm 0.11	1.19 \pm 0.05	1.47 \pm 0.08	73.12 \pm 2.52	39.12 \pm 2.04
Sham Week 2, n=10	3.96 \pm 0.17	2.53 \pm 0.11	1.14 \pm 0.04	1.43 \pm 0.07	73.33 \pm 2.21	38.01 \pm 1.58
MI-only Week 2 post injury, n=8	5.31 \pm 0.23	4.08 \pm 0.22*	0.85 \pm 0.05	0.91 \pm 0.09	35.58 \pm 3.57	16.48 \pm 1.67
MI+Patch Week 2 post injury, n=8	4.63 \pm 0.08* [●]	3.56 \pm 0.16* [●]	1.11 \pm 0.05 [●]	1.36 \pm 0.09 [●]	48.17 \pm 3.52* [●]	21.69 \pm 2.17* [●]
MI+Patch+CM Week 2 post injury, n=8	4.45 \pm 0.16* [●]	3.10 \pm 0.15* [●]	1.12 \pm 0.04* [●]	1.40 \pm 0.06* [●]	64.18 \pm 1.90* ^{●■}	30.35 \pm 1.30* ^{●■}
MI+Patch+FSTL1 Week 2 post injury, n=9	4.77 \pm 0.23* [●]	3.72 \pm 0.33* [●]	1.02 \pm 0.07* [●]	1.14 \pm 0.08* [●]	50.31 \pm 5.64* ^{●▲}	23.17 \pm 3.05* ^{●▲}
Sham Week 4, n=10	3.98 \pm 0.09	2.55 \pm 0.12	1.02 \pm 0.07	1.41 \pm 0.07	71.32 \pm 2.63	35.17 \pm 1.60
MI-only Week 4 post injury, n=8	5.27 \pm 0.20*	4.55 \pm 0.24*	0.70 \pm 0.04*	0.78 \pm 0.06*	35.32 \pm 2.64*	12.71 \pm 1.28*
MI+Patch Week 4 post injury, n=8	5.04 \pm 0.12* [●]	3.77 \pm 0.11* [●]	1.02 \pm 0.08 [●]	1.28 \pm 0.08 [●]	51.43 \pm 1.62* [●]	22.71 \pm 0.92* [●]
MI+Patch+CM Week 4 post injury, n=8	4.35 \pm 0.18 [■]	3.20 \pm 0.24* ^{●■}	1.23 \pm 0.05 [●]	1.53 \pm 0.06 [●]	62.15 \pm 4.36 [■]	27.52 \pm 2.50* ^{●■}
MI+Patch+FSTL1 Week 4 post injury, n=9	4.24 \pm 0.12 [■]	2.86 \pm 0.13* [●]	1.15 \pm 0.06 [●]	1.50 \pm 0.08 [●]	67.73 \pm 1.79 [■]	32.87 \pm 1.26* ^{●■}

The patch was implanted simultaneously with injury. * $P < 0.05$ statistically significant difference in comparison with Sham. Black circle ($P < 0.05$), statistically significant difference in comparison with MI-only. black square ($P < 0.05$) statistically significant difference in comparison with MI plus patch; black triangle statistically significant difference ($P < 0.05$) in comparison with MI plus patch plus CM.

Extended Data Table 2 | Raw echocardiography values (average \pm s.e.m.) in a long term (months 2 and 3) post treatment, in a mouse model of permanent LAD ligation

	LVIDd (mm)	LVIDs (mm)	LVPWd (mm)	LVPWs (mm)	EF (%)	FS (%)
Sham Month 2, n=4	4.26 \pm 0.05	2.68 \pm 0.03	1.20 \pm 0.03	1.41 \pm 0.07	72.00 \pm 0.41	35.75 \pm 0.25
MI Month 2, n=4	5.53 \pm 0.17*	4.76 \pm 0.19*	0.73 \pm 0.04*	0.91 \pm 0.09*	33.98 \pm 3.24*	13.88 \pm 1.56*
MI+Patch Month 2, n=4	5.05 \pm 0.09* [●]	3.92 \pm 0.04* [●]	0.92 \pm 0.09*	1.27 \pm 0.08*	51.13 \pm 2.70* [●]	21.50 \pm 1.44* [●]
MI+Patch+FSTL1 Month 2, n=4	3.96 \pm 0.16 [■]	2.58 \pm 0.09 [■]	1.27 \pm 0.04 [■]	1.50 \pm 0.09 [■]	68.25 \pm 0.75 [■]	33.19 \pm 0.70 [■]
Sham Month 3, n=4	4.41 \pm 0.13	3.01 \pm 0.13	1.12 \pm 0.02	1.34 \pm 0.01	67.00 \pm 1.74	32.38 \pm 1.28
MI Month 3, n=4	4.93 \pm 0.27*	4.14 \pm 0.16*	0.90 \pm 0.03*	0.99 \pm 0.09*	38.00 \pm 2.65*	15.33 \pm 1.20*
MI+Patch Month 3, n=4	4.94 \pm 0.23	3.68 \pm 0.22*	0.90 \pm 0.10	1.18 \pm 0.02*	53.50 \pm 2.63* [●]	23.58 \pm 1.57* [●]
MI+Patch+FSTL1 Month 3, n=4	4.65 \pm 0.04 [●]	3.19 \pm 0.06 [■]	0.93 \pm 0.09	1.17 \pm 0.07 [■]	66.19 \pm 1.12 [■]	31.38 \pm 1.46 [■]

The patch was implanted simultaneously with injury. *P < 0.05, statistically significant difference in comparison with sham. Black circle, statistically significant difference (P < 0.05) in comparison with MI-only, black square, statistically significant difference (P < 0.05) in comparison with MI plus patch; black triangle, statistically significant difference (P < 0.05) in comparison with MI plus patch plus CM.

Extended Data Table 3 | Raw echocardiography values (average \pm s.e.m) of delayed grafting in a mouse model of ischaemia/reperfusion

	LVIDd (mm)	LVIDs (mm)	LVPWd (mm)	LVPWs (mm)	EF (%)	FS (%)
Sham Baseline, n=4	3.62 \pm 0.15	2.39 \pm 0.14	1.13 \pm 0.09	1.44 \pm 0.10	74.98 \pm 2.57	35.67 \pm 1.03
IR Baseline, n=4	3.90 \pm 0.10	2.59 \pm 0.03	1.26 \pm 0.05	1.45 \pm 0.11	73.06 \pm 2.98	36.81 \pm 2.36
IR+Patch+FSTL1 Baseline, n=4	4.09 \pm 0.09	2.67 \pm 0.06	1.27 \pm 0.03	1.50 \pm 0.05	77.50 \pm 2.57	40.83 \pm 2.24
Sham Week 1 post injury Pre-graft, n=4	3.62 \pm 0.07	2.27 \pm 0.05	1.16 \pm 0.05	1.43 \pm 0.03	71.67 \pm 3.33	35.50 \pm 2.50
IR Week 1 post injury Pre-graft, n=4	4.27 \pm 0.16*	3.33 \pm 0.16*	1.00 \pm 0.03*	1.12 \pm 0.03	48.33 \pm 0.51*	22.13 \pm 0.84*
IR+Patch+FSTL1 Week 1 post injury Pre-graft, n=4	4.34 \pm 0.32*	3.49 \pm 0.32*	1.02 \pm 0.03*	0.99 \pm 0.05	46.69 \pm 3.04*	19.74 \pm 1.57*
Sham Week 3 post injury Week 2 post graft, n=4	3.78 \pm 0.24	2.46 \pm 0.15	1.15 \pm 0.06	1.35 \pm 0.06	69.95 \pm 3.59	34.04 \pm 2.44
IR Week 3 post injury Week 2 post graft, n=3	4.49 \pm 0.16*	3.56 \pm 0.14*	0.91 \pm 0.07*	0.99 \pm 0.04*	47.22 \pm 2.54*	20.12 \pm 1.36*
IR+Patch+FSTL1 Week 3 post injury Week 2 post graft, n=4	4.34 \pm 0.12	2.77 \pm 0.14 [•]	1.08 \pm 0.00* [•]	1.44 \pm 0.04 [•]	70.78 \pm 2.42 [•]	33.67 \pm 2.19* [•]
Sham Week 5 post injury Week 4 post graft, n=4	4.03 \pm 0.16	2.75 \pm 0.21	1.05 \pm 0.09	1.34 \pm 0.10	66.46 \pm 3.92	33.08 \pm 2.47
IR Week 5 post injury Week 4 post graft, n=3	4.63 \pm 0.16	3.84 \pm 0.10*	0.87 \pm 0.05*	0.95 \pm 0.09*	42.08 \pm 0.96*	16.81 \pm 1.13*
IR+Patch+FSTL1 Week 5 post injury Week 4 post graft, n=4	4.09 \pm 0.15*	2.57 \pm 0.12 [•]	1.20 \pm 0.04 [•]	1.41 \pm 0.10 [•]	73.16 \pm 1.13 [•]	34.44 \pm 0.87 [•]

Data obtained at baseline (pre-injury, pre-grafting), weeks 1 (post injury, pre-grafting) and weeks 3, 5 (post-injury, post-grafting). The patch was implanted 1 week after injury. *P < 0.05, statistically significant difference in comparison with sham; black circle indicates statistically significant difference (P < 0.05) in comparison with MI-only; black square indicates statistically significant difference (P < 0.05) in comparison with MI plus patch; black triangle statistically significant difference (P < 0.05) in comparison with MI plus patch plus CM.

Structure of the toxic core of α -synuclein from invisible crystals

Jose A. Rodriguez^{1*}, Magdalena I. Ivanova^{1*†}, Michael R. Sawaya^{1*}, Duilio Cascio^{1*}, Francis E. Reyes^{2*}, Dan Shi², Smriti Sangwan¹, Elizabeth L. Guenther¹, Lisa M. Johnson¹, Meng Zhang¹, Lin Jiang^{1†}, Mark A. Arbing¹, Brent L. Nannenga², Johan Hattne², Julian Whitelegge³, Aaron S. Brewster⁴, Marc Messerschmidt^{5†}, Sébastien Boutet⁵, Nicholas K. Sauter⁴, Tamir Gonen² & David S. Eisenberg¹

The protein α -synuclein is the main component of Lewy bodies, the neuron-associated aggregates seen in Parkinson disease and other neurodegenerative pathologies. An 11-residue segment, which we term NACore, appears to be responsible for amyloid formation and cytotoxicity of human α -synuclein. Here we describe crystals of NACore that have dimensions smaller than the wavelength of visible light and thus are invisible by optical microscopy. As the crystals are thousands of times too small for structure determination by synchrotron X-ray diffraction, we use micro-electron diffraction to determine the structure at atomic resolution. The 1.4 Å resolution structure demonstrates that this method can determine previously unknown protein structures and here yields, to our knowledge, the highest resolution achieved by any cryo-electron microscopy method to date. The structure exhibits protofibrils built of pairs of face-to-face β -sheets. X-ray fibre diffraction patterns show the similarity of NACore to toxic fibrils of full-length α -synuclein. The NACore structure, together with that of a second segment, inspires a model for most of the ordered portion of the toxic, full-length α -synuclein fibril, presenting opportunities for the design of inhibitors of α -synuclein fibrils.

The presynaptic protein α -synuclein, found in both soluble and membrane-associated fractions of the brain, aggregates in Parkinson disease (PD). These aggregates are the main component of Lewy bodies, the defining histological feature of this neurodegenerative disease, and have been shown to accompany neuronal damage¹. Two other observations point to aggregated α -synuclein as a molecular cause of PD². The first is that families with inherited forms of PD carry mutations in α -synuclein, such as A53T, and abundant Lewy bodies^{3–5}. The second is that families with duplicated or triplicated genes encoding α -synuclein develop early-onset PD, presumably because at high local concentrations α -synuclein is forced into amyloid^{6,7}.

Our focus is on a central segment of α -synuclein, residues 68–78, that we term NACore (Fig. 1), because of its critical role in both the aggregation and cytotoxicity of α -synuclein. NACore lies within a 35-residue domain of α -synuclein termed NAC (non-amyloid- β component, originally reported to be deposited with amyloid- β in the brains of Alzheimer's disease patients), which has been established as necessary and sufficient for the aggregation and toxicity of α -synuclein^{8–12} (Extended Data Fig. 1). For example, deletion of residues 71–82 prevents aggregation of α -synuclein *in vitro*, and abolishes both its aggregation and neurotoxicity in a *Drosophila* model of PD¹². Yet this segment in isolation from the rest of α -synuclein readily forms amyloid fibrils and is highly cytotoxic^{13,14}. Also, β -synuclein, the close homologue of α -synuclein, which does not aggregate and is not found in Lewy bodies, differs in sequence from α -synuclein principally by the lack of residues 74–84 that are part of NACore⁹.

Segments outside NAC also influence the aggregation of α -synuclein and have been associated with fibril structure^{15–17}. In brain extracts from

patients with multiple system atrophy, the core of α -synuclein fibrils extends approximately from residue 30 to 100¹⁸. Also the A53T mutation of α -synuclein can accelerate its transition into the amyloid state, and hence accelerate PD¹⁹. This mutation was found to induce the onset of PD at an early age²⁰, and consistent with this, α -synuclein containing this A53T mutation forms fibrils *in vitro* more rapidly than the wild type³. Thus we carried out screens for crystals of peptide segments within the NAC domain and adjacent regions, seeking structural information on the molecular basis of aggregation and toxicity of α -synuclein.

Structure determination by MicroED

Extensive crystal screens of two segments, NACore, residues 68GAVVTGVTAVA₇₈, and PreNAC, 47GVVHGVTVA₅₆, seemingly produced non-crystalline, amorphous aggregates. But upon examination using electron microscopy, we found the aggregates to be clusters of elongated nanocrystals only 50–300 nm in cross section and thus invisible by conventional light microscopy (Fig. 1). We confirmed well-ordered crystallinity of NACore at both the SACLA and LCLS free electron lasers. We also found that a nine-residue fragment within the NACore, which we term SubNACore, 69AVVTGVTAV₇₇, yielded crystals 1,000–10,000 times larger in volume than the NACore nanocrystals (Fig. 1). We were therefore able to apply synchrotron methods^{21,22} to these larger crystals to determine the structure of their amyloid-like fibrils. Although this nine-residue fragment is missing only two residues compared with NACore, it is not as toxic²³, offering some insight into the toxicity of α -synuclein, as described below.

To determine the structure of the invisible crystals of NACore and PreNAC, we turned to micro-electron diffraction (MicroED)^{24–26}. In

¹Howard Hughes Medical Institute, UCLA-DOE Institute, Departments of Biological Chemistry and Chemistry and Biochemistry, Box 951570, UCLA, Los Angeles, California 90095-1570, USA. ²Howard Hughes Medical Institute, Janelia Research Campus, 19700 Helix Drive, Ashburn, Virginia 20147, USA. ³Box 42, NPI-Semel Institute, 760 Westwood Plaza, UCLA, Los Angeles, California 90024, USA.

⁴Physical Biosciences Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720, USA. ⁵Linac Coherent Light Source, SLAC National Accelerator Laboratory, Menlo Park, California 94025, USA. [†]Present addresses: Department of Neurology and Program of Biophysics, University of Michigan School of Medicine, Ann Arbor, Michigan 48109, USA (M.I.); Department of Neurology, UCLA, Los Angeles, California 90095, USA (L.J.); National Science Foundation BioXFEL Science and Technology Center, Buffalo, New York 14203, USA (M.M.).

*These authors contributed equally to this work.

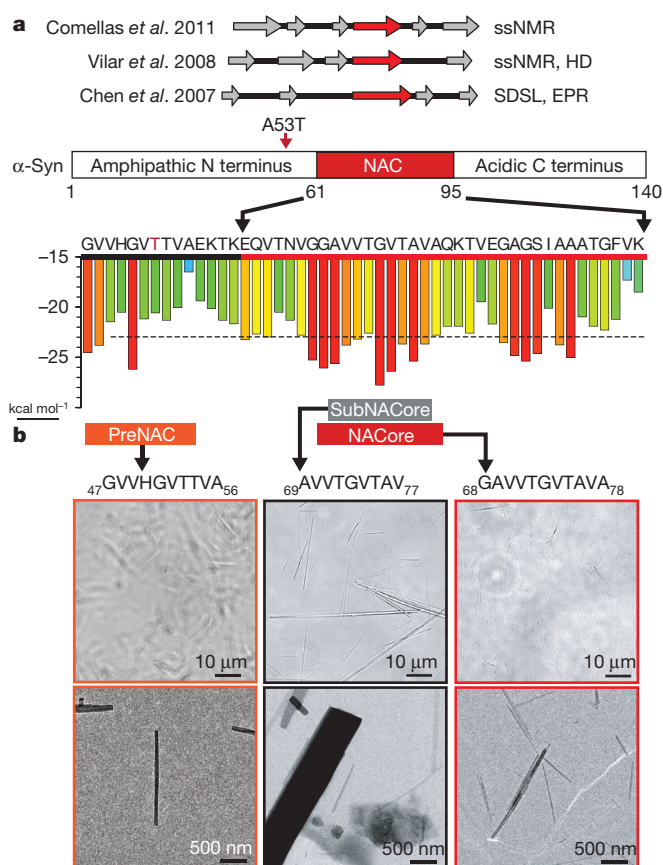


Figure 1 | NACore (residues 68–78) is the fibril-forming core of the NAC domain of full-length α -synuclein. **a**, Top, segments identified as β -strands by electron paramagnetic resonance (EPR), solid state nuclear magnetic resonance (ssNMR), hydrogen-deuterium exchange (HD), and site directed spin labelling (SDSL)^{17,40,41}. Bottom, predictions of the propensity of six-residue segments to form amyloid fibrils. The vertical axis indicates the propensity of steric zipper formation in Rosetta energy units⁴². Zipper-forming segments are predicted where bars cross the $-23 \text{ kcal mol}^{-1}$ threshold marked by the dashed line. Blue-to-red color gradient indicates weak-to-strong propensity of steric zipper formation. The A53T early-onset Parkinson mutation is indicated by a red arrow and red letter T. **b**, The miniscule size of the preNAC and NACore crystals used for MicroED is illustrated by this comparison to SubNACore microcrystals. Scale comparisons are illustrated on two magnifications using phase contrast light microscope images and electron micrographs, in which individual NACore and PreNAC nanocrystals are indistinguishable by light microscopy.

MicroED, an extremely low-dose electron beam is directed onto a nanocrystal within a transmission electron microscope under cryogenic conditions, yielding diffraction patterns such as those in Fig. 2. As the wavelength used in our experiments at 200 keV is very small (0.025 Å), the Ewald sphere is essentially flat, resulting in diffraction patterns that closely resemble a 2D slice through 3D reciprocal space. As the crystal is continuously rotated in the beam, a series of such diffraction patterns is collected²⁵. Scaling together diffraction data collected from multiple crystals produces a full 3D diffraction data set. MicroED has been successfully applied to the well-known structures of hen egg-white lysozyme^{25,26}, bovine liver catalase²⁷ and Ca^{2+} -ATPase²⁸. But NACore and PreNAC are the first previously unknown structures determined by MicroED.

For NACore and PreNAC, we collected MicroED patterns from nano-crystals that lay preferentially oriented, flat on the surface of a holey carbon Quantifoil grid, in a frozen-hydrated state. Grids were first screened for appropriately sized crystals, and candidate crystals screened for diffraction. We used crystals showing strong diffraction for data collection by continuous unidirectional rotation about a fixed

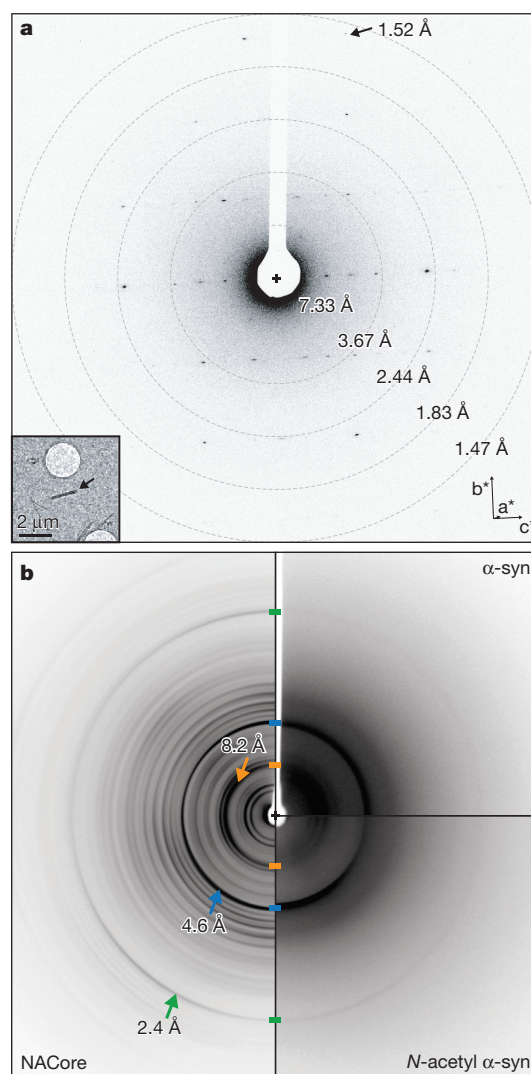


Figure 2 | Diffraction from NACore nanocrystals is similar to that from full length α -synuclein fibrils. **a**, Single-crystal electron diffraction pattern obtained during MicroED data collection (see text). Equally spaced concentric rings denote resolution shells. The highest resolution spot is at 1.52 Å (arrow). The inset shows the over-focused image of the diffracting crystal (arrow), which is $\sim 1,480 \times 200 \times 200 \text{ nm}$. Orientation of the reciprocal cell axes are indicated by the arrows labelled a^* , b^* , c^* . **b**, Composite of fibril diffraction patterns from α -synuclein (α -syn) preparations and NACore. Full-length α -synuclein reveals reflections that match those from NACore and N-terminally acetylated α -synuclein. The two patterns of full-length α -synuclein share with NACore three major peaks denoted by arrows: 8.2 Å (orange), 4.6 Å (blue), and 2.4 Å (green). The origin of these peaks can be traced to the (0,0,2), [(1,1,1), (-1,1,1)], and (0,2,0) planes in the NACore structure, respectively. We attribute the strong 8.2 Å reflection to the spacing between adjacent pairs of β -sheets.

axis, acquiring a series of diffraction frames at fixed time intervals²⁵. The needle-shaped crystals typically exceeded the length needed for MicroED; those that were unbent and 100 to 300 nm wide produced the best diffraction patterns. Data from multiple crystals were integrated, scaled and merged together (Extended Data Table 1).

The multi-crystal NACore and PreNAC data sets were phased by molecular replacement, using the atomic model of SubNACore and an ideal β -strand model, respectively, as probes. Residues of NACore which were missing from the SubNACore probe were clearly revealed in a difference density map calculated from NACore observed structure factor amplitudes and phases from the SubNACore probe structure (Extended Data Fig. 2). After subsequent refinement, two water

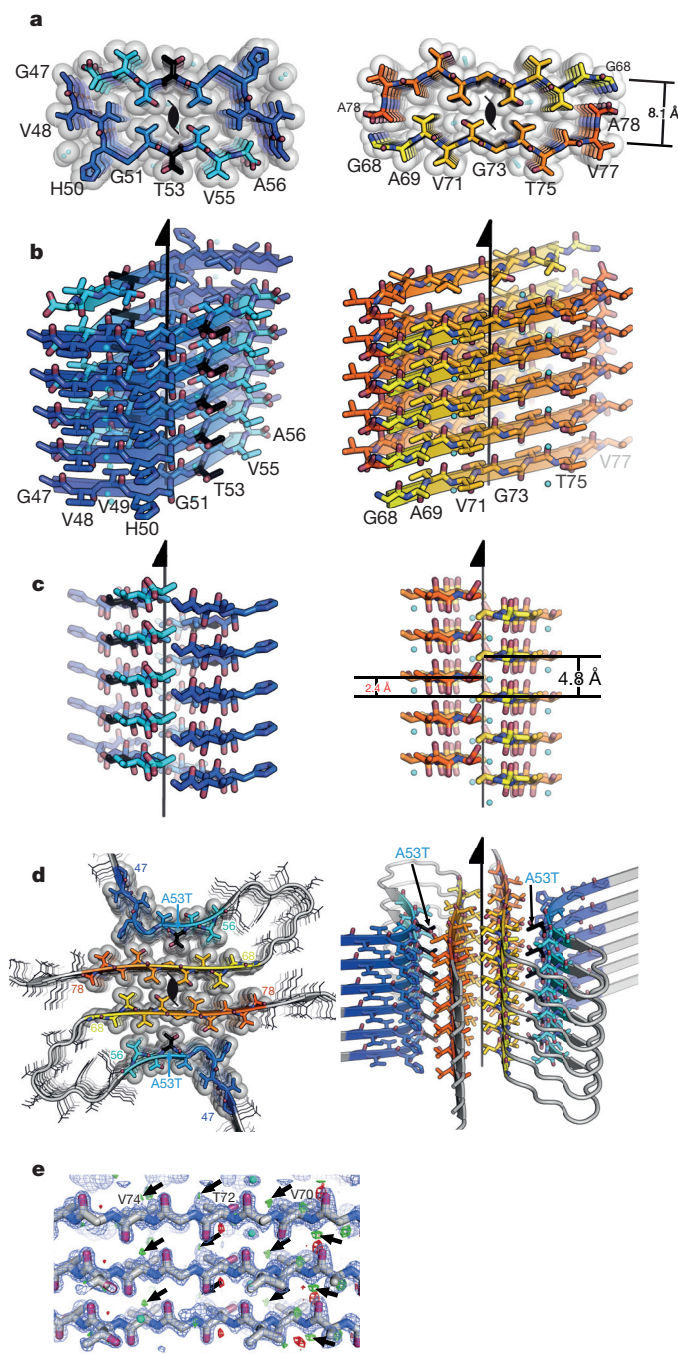


Figure 3 | Structure of the amyloid core of α -synuclein. **a**, The crystal structure of NACore (orange) reveals pairs of sheets as in the spines of amyloid fibrils. The A53T mutation (black) is shown in PreNAC (blue). The sheets in both structures are related by the 2_1 fibril axes shown in black. The gaps left by the interface are filled with water molecules which hydrogen-bond to the threonine residues (partially showing aqua spheres). **b**, **c**, Orthogonal views of the fibrillar assemblies. **d**, A speculative model of an α -synuclein protofibril containing the A53T mutation (black), where the strong interface of NACore (orange) forms the core of the fibril and its weaker interface interacts with PreNAC (blue). **e**, The locations of 5 out of a possible 73 protons are suggested by small, positive $F_o - F_c$ density (green contoured at 2.8σ , shown by arrows). The blue mesh is $2F_o - F_c$ density contoured at 1.4σ .

molecules, and several hydrogen atoms, were observed (Fig. 3e). Full models of NACore and PreNAC were refined against the MicroED data, producing structures at 1.4 \AA resolution with acceptable R factors (Extended Data Table 1). Electron scattering factors were used in the refinement calculations²⁹.

NACore structure

The structure of the NACore peptide chain is a nearly fully extended β -strand (Fig. 3 and Extended Data Fig. 3). These NACore strands stack in-register into β -sheets, as had been predicted by site-directed spin labelling^{16,17}. The sheets are paired (Fig. 3b), as is usual in amyloid spines, and the pairs of sheets form typical steric-zipper protofilaments (Fig. 3c), previously seen as the spines in many amyloid-like fibrils formed from short segments of fibril-forming proteins²¹. The unusual features of this steric zipper are that the 11-residue width of the zipper is longer than has been previously observed²², and each pair of sheets contains two water molecules, each associated with a threonine side chain within the interface. Most steric zippers are completely dry. Also, in our crystals of NACore, each sheet forms two snug interfaces: interface A, with 268 \AA^2 of buried accessible surface area per chain, is more extensive and presumably stronger than interface B (167 \AA^2), because the terminal residues of the chains in opposing sheets bend towards each other (Fig. 3 and Extended Data Fig. 4). The structure of PreNAC reveals a peptide chain that forms a β -strand kinked at Gly51. These strands are arranged into pairs of β -sheets that, like the NACore structure, interdigitate to form steric zipper protofilaments (Fig. 3). Of special note, a five-residue segment of PreNAC (₅₁GVTTV₅₅) differs in only one residue from a five-residue segment of NACore (₇₃GVTAV₇₇), and their α -carbons superimpose closely with a root mean square deviation (r.m.s.d.) of 1.5 \AA (Extended Data Fig. 4). This means that the weaker interface B of NACore mimics a hypothetical interface between NACore and PreNAC (Fig. 3d).

Relevance of NACore to Parkinson disease

The relationship of the structure of NACore to fibrils of full length α -synuclein is established by the resemblance of their diffraction patterns. Specifically, the fibre diffraction pattern of aligned fibrils of full-length and N-terminally acetylated³⁰ α -synuclein protein display the same principal peaks as the diffraction of aligned NACore nanocrystals (Fig. 2). All three fibrils display the strong reflection at 2.4 \AA in their diffraction patterns. As seen in Fig. 3 and Extended Data Fig. 5 this reflection arises in NACore because one β -sheet of the steric zipper is translated along the fibre axis with respect to the other β -sheet by 2.4 \AA , one half the 4.8 \AA spacing between β -strands, permitting the two sheets to interdigitate tightly together. All three share a strong 4.6 \AA reflection, which in NACore results from both the stacking of β -strands and the staggering between adjacent β -sheets of the steric zipper, while a shared reflection at near 8.2 \AA probably arises from the distance between the adjacent pairs of β -sheets that make up the α -synuclein fibril (Fig. 2 and Extended Data Fig. 5). This comparison of fibre diffraction patterns (Extended Data Table 2) strongly suggests that the structure of NACore is similar to the spine of the toxic fibrils of full α -synuclein.

The combined structures of NACore and PreNAC allow us to construct a speculative model for much of the ordered segments of the A53T early-onset mutant α -synuclein (Fig. 3d). Experimental support of this model comes from the agreement of its simulated fibre diffraction with the measured diffraction patterns of α -synuclein and N-acetyl α -synuclein fibrils, as well as aligned NACore nanocrystals (Extended Data Table 2). Above we hypothesized that the weaker interface B of NACore might mimic an intramolecular interaction of PreNAC with NACore (Fig. 3). In fact, the interacting side chains in the weaker NACore interface B (G73, T75 and V77) are identical to the side chains (G51, T53, V55) interacting in the hypothetical interface of PreNAC with NACore. Assuming that this interface actually forms in fibrils of the early-onset mutant A53T, we built the model shown in Fig. 3d. The hypothetical interface of this model offers a possible reason for a greater propensity of the A53T mutant to aggregate than the wild-type sequence, conceivably leading to the early onset of PD.

The identity and structure of the cytotoxic amyloid formed by α -synuclein remains a subject of intensive research^{19,31–35}. The weight

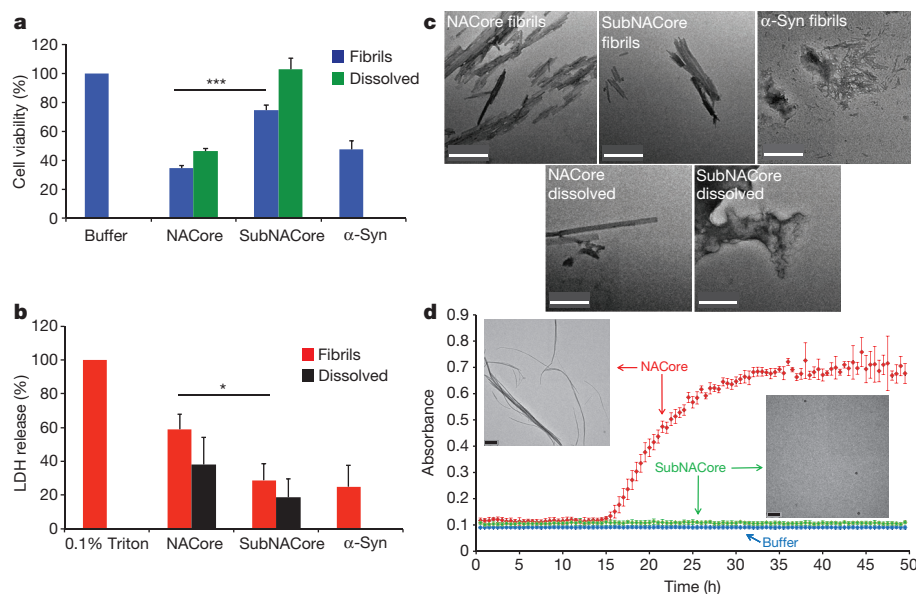


Figure 4 | NACore aggregates faster than SubNACore and is more cytotoxic to cultured cells. **a**, **b**, Cytotoxicity of NACore, SubNACore and α -synuclein measured on PC12 cells using a 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT) assay (**a**) and a lactate dehydrogenase (LDH)-release assay (**b**). In both assays NACore is more toxic than SubNACore. Also, shaken fibrils are more toxic than an equal concentration of freshly dissolved sample. Results shown as mean \pm s.e.m. based on triplicate samples. A *t*-test was used to measure statistical significance; **P* < 0.05, ****P* < 0.001. **c**, Representative electron micrographs of NACore, SubNACore

and α -synuclein (α -syn) samples tested for cytotoxicity. NACore and α -synuclein show abundant fibrils but SubNACore shows few. NACore also forms fibrils immediately upon dissolving, whereas SubNACore shows no fibres, but instead amorphous aggregates. Scale bar, 500 nm. **d**, NACore and SubNACore were aggregated in identical conditions and monitored by turbidity. NACore begins to aggregate after 15 h while SubNACore forms no aggregates for up to 50 h. Electron microscopy of the samples at 50 h confirmed the turbidity readings (insets; scale bars, 2 μ m), with error bars denoting standard deviation based on triplicate samples.

of evidence over the past decade has tilted scientific opinion from the fully developed amyloid fibrils found in Lewy bodies as the toxic entities to smaller, transient amyloid oligomers. Yet recently, quantitative arguments have been put forward in favour of fibrils³⁶. Our experiments of the cytotoxicity of NACore on PC12 cells (Fig. 4) are consistent with the view that fibrils are toxic: we find that NACore shaken and aggregated for 72 h displays abundant fibrils, is more toxic than freshly dissolved NACore (Fig. 4), and is comparably toxic to similarly aggregated full α -synuclein. We also find greater cytotoxicity of NACore than SubNACore, which is shorter by two residues. This is consistent with the more rapid fibril formation of NACore than of SubNACore (Fig. 4d). These observations do not rule out the formation of a non-fibrillar, oligomeric assembly, present, but undetected, in our aggregated samples of NACore and α -synuclein. Of course, NACore is merely a fragment of full length α -synuclein, and lacks most of the membrane-binding motifs of the N terminus of the protein, which have been implicated in membrane disruption^{37,38}. Yet it is clear that NACore is the minimum entity that recapitulates all the features of full length α -synuclein aggregation and toxicity.

MicroED diffraction of invisible crystals

The miniscule size of NACore crystals is typical of amyloid and also of various other biological crystals of interest. For amyloid crystals, our speculation is that the tiny size is a consequence of the natural twist of β -sheets that form the protofilaments of the fibrils. The crystal lattice restrains the twist, creating a strain in these crystals, which increases as crystals grow. Eventually this strain prevents further addition of β -strands, limiting the thickness of the needle crystals. In our experience, longer segments (for example, 11 residues compared to 9 residues) limit crystal growth even more; in the case of 11-residue NACore and 10-residue PreNAC, the strain produces nanocrystals, invisible by optical microscopy. These crystals are too small for mounting and conventional synchrotron data collection, but are ideally suited for analysis by MicroED. They are $\sim 10^{10}$ times smaller

than Perutz's haemoglobin crystals and $\sim 10^{12}$ times smaller than von Laue's CuSO_4 crystal, which yielded the first X-ray diffraction pattern. Our structures of NACore and PreNAC demonstrate that MicroED is capable of determining new and accurate structures of biological material at atomic resolutions. This finding paves the way for applications of MicroED to other biological substances of importance for which only nanocrystals can be grown. In our particular application, we have been able to learn the atomic arrangement of the core of the crucial NAC domain. This presents opportunities for structure-based design of inhibitors of amyloid formation of α -synuclein³⁹.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 30 October 2014; accepted 13 August 2015.

Published online 9 September 2015.

- Spillantini, M. G. *et al.* α -Synuclein in Lewy bodies. *Nature* **388**, 839–840 (1997).
- Goedert, M., Spillantini, M. G., Del Tredici, K. & Braak, H. 100 years of Lewy pathology. *Nature Rev. Neurol.* **9**, 13–24 (2013).
- Polymeropoulos, M. H. *et al.* Mutation in the α -synuclein gene identified in families with Parkinson's disease. *Science* **276**, 2045–2047 (1997).
- Krüger, R. *et al.* Ala30Pro mutation in the gene encoding α -synuclein in Parkinson's disease. *Nature Genet.* **18**, 106–108 (1998).
- Zarranz, J. J. *et al.* The new mutation, E46K, of α -synuclein causes Parkinson and Lewy body dementia. *Ann. Neurol.* **55**, 164–173 (2004).
- Ibáñez, P. *et al.* Causal relation between α -synuclein gene duplication and familial Parkinson's disease. *Lancet* **364**, 1169–1171 (2004).
- Singleton, A. B. *et al.* α -Synuclein locus triplication causes Parkinson's disease. *Science* **302**, 841 (2003).
- Ueda, K. *et al.* Molecular cloning of cDNA encoding an unrecognized component of amyloid in Alzheimer disease. *Proc. Natl Acad. Sci. USA* **90**, 11282–11286 (1993).
- Biere, A. L. *et al.* Parkinson's disease-associated α -synuclein is more fibrillogenic than β - and γ -synuclein and cannot cross-seed its homologs. *J. Biol. Chem.* **275**, 34574–34579 (2000).
- Giasson, B. I., Murray, I. V. J., Trojanowski, J. Q. & Lee, V. M.-Y. A hydrophobic stretch of 12 amino acid residues in the middle of α -synuclein is essential for filament assembly. *J. Biol. Chem.* **276**, 2380–2386 (2001).
- Du, H.-N. *et al.* A peptide motif consisting of glycine, alanine, and valine is required for the fibrillization and cytotoxicity of human α -synuclein. *Biochemistry* **42**, 8870–8878 (2003).

12. Periquet, M., Fulga, T., Myllykangas, L., Schlossmacher, M. G. & Feany, M. B. Aggregated α -synuclein mediates dopaminergic neurotoxicity *in vivo*. *J. Neurosci.* **27**, 3338–3346 (2007).
13. Han, H., Weinreb, P. H. & Lansbury, P. T. The core Alzheimer's peptide NAC forms amyloid fibrils, which seed and are seeded by β -amyloid: is NAC a common trigger or target in neurodegenerative disease? *Chem. Biol.* **2**, 163–169 (1995).
14. El-Agnaf, O. M. *et al.* Aggregates from mutant and wild-type α -synuclein proteins and NAC peptide induce apoptotic cell death in human neuroblastoma cells by formation of β -sheet and amyloid-like filaments. *FEBS Lett.* **440**, 71–75 (1998).
15. Crowther, R. A., Daniel, S. E. & Goedert, M. Characterisation of isolated α -synuclein filaments from substantia nigra of Parkinson's disease brain. *Neurosci. Lett.* **292**, 128–130 (2000).
16. Der-Sarkissian, A., Jao, C. C., Chen, J. & Langen, R. Structural organization of α -synuclein fibrils studied by site-directed spin labeling. *J. Biol. Chem.* **278**, 37530–37535 (2003).
17. Chen, M., Margittai, M., Chen, J. & Langen, R. Investigation of α -synuclein fibril structure by site-directed spin labeling. *J. Biol. Chem.* **282**, 24970–24979 (2007).
18. Miake, H., Mizusawa, H., Iwatsubo, T. & Hasegawa, M. Biochemical characterization of the core structure of α -synuclein filaments. *J. Biol. Chem.* **277**, 19213–19219 (2002).
19. Conway, K. A., Harper, J. D. & Lansbury, P. T. Accelerated *in vitro* fibril formation by a mutant α -synuclein linked to early-onset Parkinson disease. *Nature Med.* **4**, 1318–1320 (1998).
20. Polymeropoulos, M. H. *et al.* Mutation in the α -synuclein gene identified in families with Parkinson's disease. *Science* **276**, 2045–2047 (1997).
21. Nelson, R. *et al.* Structure of the cross- β spine of amyloid-like fibrils. *Nature* **435**, 773–778 (2005).
22. Sawaya, M. R. *et al.* Atomic structures of amyloid cross- β spines reveal varied steric zippers. *Nature* **447**, 453–457 (2007).
23. Bodles, A. M., Guthrie, D. J., Greer, B. & Irvine, G. B. Identification of the region of non-A β component (NAC) of Alzheimer's disease amyloid responsible for its aggregation and toxicity. *J. Neurochem.* **78**, 384–395 (2001).
24. Nannenga, B. L. & Gonen, T. Protein structure determination by MicroED. *Curr. Opin. Struct. Biol.* **27**, 24–31 (2014).
25. Nannenga, B. L., Shi, D., Leslie, A. G. W. & Gonen, T. High-resolution structure determination by continuous-rotation data collection in MicroED. *Nature Methods* **11**, 927–930 (2014).
26. Shi, D., Nannenga, B. L., Iadanza, M. G. & Gonen, T. Three-dimensional electron crystallography of protein microcrystals. *eLife* **2**, e01345 (2013).
27. Nannenga, B. L., Shi, D., Hattne, J., Reyes, F. E. & Gonen, T. Structure of catalase determined by MicroED. *eLife* **3**, e03600 (2014).
28. Yonekura, K., Kato, K., Ogasawara, M., Tomita, M. & Toyoshima, C. Electron crystallography of ultrathin 3D protein crystals: atomic model with charges. *Proc. Natl Acad. Sci. USA* **112**, 3368–3373 (2015).
29. Doyle, P. A. & Turner, P. S. Relativistic Hartree–Fock X-ray and electron scattering factors. *Acta Crystallogr. A* **24**, 390–397 (1968).
30. Sarafian, T. A. *et al.* Impairment of mitochondria in adult mouse brain overexpressing predominantly full-length, N-terminally acetylated human α -synuclein. *PLoS ONE* **8**, e63557 (2013).
31. Caughey, B. & Lansbury, P. T. Protofibrils, pores, fibrils, and neurodegeneration: separating the responsible protein aggregates from the innocent bystanders. *Annu. Rev. Neurosci.* **26**, 267–298 (2003).
32. Danzer, K. M., Schnack, C., Sutcliffe, A., Hengerer, B. & Gillardon, F. Functional protein kinase arrays reveal inhibition of p-21-activated kinase 4 by α -synuclein oligomers. *J. Neurochem.* **103**, 2401–2407 (2007).
33. Karpinar, D. P. *et al.* Pre-fibrillar α -synuclein variants with impaired β -structure increase neurotoxicity in Parkinson's disease models. *EMBO J.* **28**, 3256–3268 (2009).
34. Winner, B. *et al.* *In vivo* demonstration that α -synuclein oligomers are toxic. *Proc. Natl Acad. Sci. USA* **108**, 4194–4199 (2011).
35. Chen, S. W. *et al.* Structural characterization of toxic oligomers that are kinetically trapped during α -synuclein fibril formation. *Proc. Natl Acad. Sci. USA* **112**, E1994–E2003 (2015).
36. Bousset, L. *et al.* Structural and functional characterization of two α -synuclein strains. *Nature Commun.* **4**, 2575 (2013).
37. Auluck, P. K., Caraveo, G. & Lindquist, S. α -Synuclein: membrane interactions and toxicity in Parkinson's disease. *Annu. Rev. Cell Dev. Biol.* **26**, 211–233 (2010).
38. Lee, J. C., Langen, R., Hummel, P. A., Gray, H. B. & Winkler, J. R. α -Synuclein structures from fluorescence energy-transfer kinetics: implications for the role of the protein in Parkinson's disease. *Proc. Natl Acad. Sci. USA* **101**, 16466–16471 (2004).
39. Sievers, S. A. *et al.* Structure-based design of non-natural amino-acid inhibitors of amyloid fibril formation. *Nature* **475**, 96–100 (2011).
40. Comellas, G. *et al.* Structured regions of α -synuclein fibrils include the early-onset Parkinson's disease mutation sites. *J. Mol. Biol.* **411**, 881–895 (2011).
41. Vilar, M. *et al.* The fold of α -synuclein fibrils. *Proc. Natl Acad. Sci. USA* **105**, 8637–8642 (2008).
42. Goldschmidt, L., Teng, P. K., Riek, R. & Eisenberg, D. Identifying the amyloids, proteins capable of forming amyloid-like fibrils. *Proc. Natl Acad. Sci. USA* **107**, 3487–3492 (2010).

Acknowledgements We thank C. Liu for supplying PC12 cells; APS staff for beam line help solving SubNACore: M. Capel, K. Rajashankar, N. Sukumar, J. Schuermann, I. Kourinov and F. Murphy at NECAT beam lines 24-ID at APS funded by the National Institute of General Medical Sciences from the National Institutes of Health (P41 GM103403) and the DOE Office of Science by Argonne National Laboratory under Contract No. DE-AC02-06CH11357. We thank the LCLS injection staff support: S. Botha, R. Shoeman and I. Schlichting. A.S.B. and N.K.S. were supported by NIH grants GM095887 and GM102520 and by the Director, Office of Science, Department of Energy (DOE) under contract DE-AC02-05CH11231 for data-processing methods. This work was supported by the US Department of Energy Office of Science, Office of Biological and Environmental Research program under award number DE-FC02-02ER63421. We also acknowledge the award MCB-0958111 from the National Science Foundation, award 1R01-AG029430 from the National Institutes of Health, award NIH-AG016570 from Alzheimer's Disease Research (ADRC) at UCLA, and HHMI for support. J.A.R. was supported by the Giannini Foundation.

Author Contributions M.I.I. characterized the α -synuclein segments and crystals. M.I.I. and S.S. conducted the toxicity assays. L.M.J. synthesized and purified NACore peptide. M.A.A. prepared the N-terminally acetylated α -synuclein. S.S. and M.Z. prepared wild-type α -synuclein. J.W. performed the mass spectrometry analyses of α -synuclein. M.I.I. and L.M.J. crystallized NACore. E.G. grew crystals of SubNACore. E.G., M.I.I. and M.R.S. collected and processed the data and solved the structure of SubNACore. L.J. and J.A.R. identified and crystallized PreNAC. J.A.R., D.S., B.L.N. and T.G. collected MicroED data on PreNAC and NACore nanocrystals. J.A.R., F.E.R., J.H., T.G., L.J., M.R.S., and D.C. processed the MicroED data and solved the structure of PreNAC and NACore. J.A.R., M.R.S., D.C., M.M. and S.B. collected XFEL diffraction from NACore nanocrystals. A.S.B. and N.K.S. processed the XFEL data. M.R.S. and L.J. built the structure model of A53T α -synuclein protofibril. J.A.R., M.I.I., M.R.S., D.C., S.S. and E.G. prepared the figures. J.A.R., M.I.I., M.R.S., D.C., T.G. and D.S.E. wrote the paper, and all authors commented on the paper.

Author Information Atomic coordinates and structure factors have been deposited in the Protein Data Bank under accession codes 4RIK (SubNACore), 4RIL (NACore) and 4ZNN (PreNAC). The maps for PreNAC and NACore have been deposited in the EMDB with accession codes EMD-3001 and EMD-3028, respectively. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to D.S.E. (david@mbi.ucla.edu) and T.G. (gonent@janelia.hhmi.org).

METHODS

Data reporting. No statistical methods were used to predetermine sample size. The experiments were not randomized. The investigators were not blinded to allocation during experiments and outcome assessment.

Crystallization. Microcrystals of SubNACore (₆₉AVVTGVTAV₇₇) were grown from synthetic peptide purchased from CS Bio. Crystals were grown at room temperature by hanging drop vaporization. Lyophilized peptide was dissolved in water at 2.9 mg ml⁻¹ concentration in 48 mM lithium hydroxide. Peptide was mixed in a 2:1 ratio with reservoir containing 0.9 M ammonium phosphate, and 0.1 M sodium acetate pH 4.6.

Nanocrystals of NACore, ₆₈GAVVTGVTAV₇₈, were grown from synthetic peptide purchased from CS Bio. Ten batches of synthesized peptide (CSBio) at a concentration of 1 mg ml⁻¹ in sterile water were shaken at 37 °C on a Torrey Pines orbital mixing plate at speed setting 9, overnight. The insoluble material was washed in 30% (w/v) glycerol then stored in water at room temperature before diffraction. The sample contained a mixture of fibrils and crystals.

Nanocrystals of PreNAC (₄₇GVVHGVTTVA₅₆) were grown from synthetic peptide purchased from InnoPep. Crystallization trials of synthesized peptide were prepared in batch. Peptide was weighed and dissolved in sterile-filtered 50 mM phosphate buffer pH 7.0 with 0.1% DMSO at a concentration of 5 mg ml⁻¹. This solution was shaken at 37 °C on a Torrey Pines orbital mixing plate at speed setting 9, overnight.

Data collection and processing. X-ray diffraction data from microcrystals of SubNACore were collected using synchrotron radiation at the Advanced Photon Source, Northeast Collaborative Access Team micro focus beam line 24-ID-E. The beam line was equipped with an ADSC Quantum 315 CCD detector. Data from a single crystal were collected in 5° wedges at a wavelength of 0.9791 Å using a 5 µm beam diameter. We used data from three different sections along the needle axis. The crystals were cryo-cooled (100 K) for data collection. Data were processed and reduced using Denzo/Scalepack from the HKL suite of programs⁴³.

Electron diffraction data from nanocrystals of NACore and PreNAC were collected using MicroED techniques^{25,26}. These nanocrystals typically clump together. To break up the clumps, an approximately 100 µl volume of nanocrystals was placed in a sonication bath for 30 min. Nanocrystals were deposited onto a Quantifoil holey-carbon EM grid in a 2–3 µl drop after appropriate dilution, which optimized for crystal density on the grid. All grids were then blotted and vitrified by plunging into liquid ethane using a Vitrobot Mark IV (FEI), then transferring to liquid nitrogen for storage. Frozen hydrated grids were transferred to a cryo-TEM using a Gatan 626 cryo-holder. Diffraction patterns and crystal images were collected using an FEG-equipped FEI Tecnai F20 TEM operating at 200 kV and recorded using a bottom mount TVIPS F416 CMOS camera with a sensor size of 4,096 × 4,096 pixels, each 15.6 × 15.6 µm. Diffraction patterns were recorded by operating the detector in rolling shutter mode with 2 × 2 pixel binning, producing a final image 2,048 × 2,048 pixels in size. Individual image frames were taken with exposure times of 3–4 s per image, using a selected area aperture with an illuminating spot size of approximately 1 µm. This geometry equates to an electron dose of less than 0.1 e⁻ per Å² per second. During each exposure, crystals were continuously rotated within the beam at a rate of 0.3° per second, corresponding to 1.2° wedge per frame. Diffraction data were collected from several crystals each oriented differently with respect to the rotation axis. These data sets each spanned wedges of reciprocal space ranging from 40° to 80°.

X-ray diffraction data from nanocrystals of NACore were collected using XFEL radiation at the CXI instrument (Coherent X-ray Imaging) at the Linear Coherent Light Source (LCLS)-SLAC. The photon energy of the X-ray pulses was 8.52 keV (1.45 Å). Each 40 fs pulse contained up to 6 × 10¹¹ photons at the sample position, taking into account a beam line transmission of 60%. The diameter of the beam was approximately 1 µm. We used a concentration of approximately 25 µl of pelleted material suspended in 1 ml water. The sample was injected into the XFEL beam using a liquid jet injector and a gas dynamic virtual nozzle⁴⁴. The micro jet width was approximately 4 µm and the flow rate was 40 µl min⁻¹. The sample caused noticeable sputtering of the liquid jet. XFEL data were processed using cctbx.xfel^{45,46}.

Calibration of the sample to detector distance in MicroED was accomplished using a polycrystalline gold standard and by referencing the prominent reflections in the electron diffraction experiment with the corresponding reflections in the XFEL data. Calibration of the *x/y* locations of the 64-tile CSPAD detector was performed by cctbx.xfel by refining the optically measured tile positions against a thermolysin data set⁴⁵.

To gain compatibility with conventional X-ray data processing programs, the MicroED diffraction images were converted from tiff or TVIPS format to the SMV crystallographic format. We used XDS to index the diffraction images⁴⁷, and XSCALE for merging and scaling together data sets originating from different

crystals. For NACore, data from four crystals were merged, while for PreNAC, data from three crystals were merged to assemble the final data sets (see Extended Data Table 1).

Structure determination. The molecular replacement solution for SubNACore was obtained using the program Phaser⁴⁸. The search model consisted of a geometrically ideal β-strand composed of nine alanine residues. Crystallographic refinements were performed with the program Refmac⁴⁹.

The molecular replacement solution for NACore was obtained using the program Phaser⁴⁸. The search model consisted of the SubNACore structure determined previously. Crystallographic refinements were performed with the program Phenix⁵⁰ and Buster⁵¹.

The molecular replacement solution for PreNAC was obtained using the program Phaser⁴⁸. The search model consisted of a geometrically ideal β-strand composed of six residues with sequence GVTVA. Crystallographic refinements were performed with the program Phenix⁵⁰ and Refmac⁴⁹.

Model building for all segments was performed using COOT⁵². Data processing and refinement statistics are reported in Extended Data Table 1. The coordinates of the final models and the structure factors have been deposited in the Protein Data Bank with PDB code 4RIK for SubNACore, 4RIL for NACore, and 4ZNN for PreNAC. The structures were illustrated using Pymol⁵³.

Protein expression and purification. The human wild-type α-synuclein construct has been previously characterized⁵⁴ (pRK172, ampicillin, T7 promoter) with sequence: MDVFMKGLSKAKEGVVAAAEKTKQGVAAEAGKTKEGVL YVGSKTKEGVVHGVATVAEKTKEQVTNVGGAVVTGVTAVAQKTEGAG SIAAATGFVKDKQLGKNEEGAPQEGILEMPVDPDNEAYEMPSEEGYQDY EPEA.

Full length α-synuclein was purified according to published protocols³⁴. The α-synuclein construct was transformed into *Escherichia coli* expression cell line BL21 (DE3) gold (Agilent Technologies) for wild-type α-synuclein protein expression. A single colony was incubated into 100 ml LB Miller broth (Fisher Scientific) supplemented with 100 µg ml⁻¹ ampicillin (Fisher Scientific) and grown overnight at 37 °C. One litre of LB (Miller) supplemented with 100 µg ml⁻¹ ampicillin in 2-l shaker flasks was incubated with 10 ml of overnight culture and grown at 37 °C until the culture reached OD₆₀₀ ≈ 0.6–0.8 as measured by a BioPhotometer UV/VIS Photometer (Eppendorf). IPTG (Isopropyl β-D-1-thiogalactopyranoside) was added to a final concentration of 0.5 mM, and grown for 4–6 h at 30 °C. Cells were harvested by centrifugation at 5,500g for 10 min at 4 °C. The cell pellet was frozen and stored at –80 °C.

The cell pellet was thawed on ice and resuspended in lysis buffer (100 mM Tris-HCl pH 8.0, 500 mM NaCl, 1 mM EDTA pH 8.0) and lysed by sonication. Crude cell lysate was clarified by centrifugation at 15,000g for 30 min at 4 °C. The clarified cell lysate was boiled and cell debris was removed by centrifugation. Protein in the supernatant was precipitated in acid at pH 3.5 through addition of HCl by titration to protein solution on ice while stirring then centrifuged for an additional 15,000g for 30 min at 4 °C. Supernatant was dialysed against buffer A (20 mM Tris-HCl, pH 8.0). After dialysis the solution was filtered through a 0.45 µm syringe (Corning) before loading onto a 20 ml HiPrep Q HP 16/10 column (GE Healthcare). The Q-HP column was washed with five column volumes of buffer A and protein eluted using a linear gradient to 100% in five column volumes of buffer B (20 mM Tris-HCl, 1 M NaCl, pH 8.0). Protein eluted at around 50–70% buffer B; peak fractions were pooled. Pooled samples were concentrated approximately tenfold using Amicon Ultra-15 centrifugal filters. Approximately 5 ml of the concentrated sample was loaded onto a HiPrep 26/60 Sephacryl S-75 HR column equilibrated with filtration buffer (25 mM sodium phosphate, 100 mM NaCl, pH 7.5). Peak fractions were pooled from the gel filtration column and dialysed against 5 mM Tris-HCl, pH 7.5, concentrated to 3 mg ml⁻¹. These were filtered through a 0.2 µm pore size filter (Corning) and stored at 4 °C.

Recombinantly expressed full-length α-synuclein with an N-terminal acetylation was prepared and purified in the following way based on a protocol detailed in ref. 16 The α-synuclein plasmid was co-expressed with a heterodimeric protein acetylation complex from *Schizosaccharomyces pombe* to acetylate the N terminus (pACYC-DUET, chloramphenicol, T7 promoter)⁵⁵. The two vectors were co-transformed into *E. coli* BL21 (DE3) using media containing both ampicillin and chloramphenicol. Cell cultures were grown in TB media containing ampicillin and chloramphenicol and induced to express α-synuclein with 0.5 mM IPTG overnight at 25 °C. Cells were harvested by centrifugation, the cell pellet then resuspended in lysis buffer (100 mM Tris-HCl pH 8.0, 500 mM NaCl, 1 mM EDTA pH 8.0, and 1 mM phenylmethylsulfonyl fluoride) and cells lysed using an Emulsiflex homogenizer (Avestin). The lysate was boiled and debris removed by centrifugation. A protein fraction was also removed by precipitation at low pH on ice followed by centrifugation. The remaining supernatant was pH adjusted by titration and dialysed against buffer A (20 mM Tris-HCl, pH 8.0,

1 mM DTT, 1 mM EDTA, pH 8.0). The resulting protein solution was loaded onto a 5 ml Q-Sepharose FF column (GE Healthcare) equilibrated with buffer A and eluted against a linear gradient of buffer B (1 M NaCl, 20 mM Tris-HCl, pH 8.0, 1 mM DTT, 1 mM EDTA, pH 8.0). Fractions containing α -synuclein were identified using SDS-PAGE, collected, concentrated and further purified by size exclusion (Sephacryl S-100 16/60, GE Healthcare) in 20 mM Tris, pH 8.0, 100 mM NaCl, 1 mM DTT, 1 mM EDTA. Purity of fractions was assessed by SDS-PAGE.

Acetylated protein was characterized by LC-MS^{30,56}. Expected average mass: 14460.1 Da for α -synuclein and 14502.1 Da for acetylated α -synuclein. Observed average mass: 14464.0 Da for α -synuclein and 14506.0 Da for acetylated α -synuclein (Extended Data Fig. 6). The shift of 4 Da between observed and expected average masses is due to instrumental error.

Fibril formation and detection. Purified α -synuclein in 50 mM Tris, 150 mM KCl, pH 7.5 was shaken at a concentration of 500 μ M at 37 °C in a Torey Pine shaker. To form the fibrillar samples of SubNACore and NACore, lyophilized peptides were dissolved to a final concentration of 500 μ M in 5 mM lithium hydroxide, 20 mM sodium phosphate pH 7.5 and 0.1 M NaCl. All samples were shaken at 37 °C in a Torey Pine shaker for 72 h. Freshly dissolved samples were prepared by dissolving lyophilized peptides immediately before addition to cells for assays.

Turbidity measurements were used to compare NACore and SubNACore aggregation. Peptide samples were freshly dissolved to 1.6 mM in a sample buffer with 5 mM LiOH and 1% DMSO and then filtered through a PVDF filter (Millipore, 0.1 μ m). Measurements were performed using a black NUNC 96 well plate with 200 μ l of sample per well (3–4 replicates per sample). The plate was agitated at 37 °C, with a 3 mm rotation diameter, at 300 r.p.m. in a Varioskan microplate reader (Thermo). Absorbance readings were recorded every 3–15 min at 340 nm.

Negative-stain transmission electron microscopy. Cytotoxicity samples were evaluated for presence of fibrils by electron microscopy. In brief, 5- μ l samples were spotted directly on freshly glow-discharged carbon-coated electron microscopy grids (Ted Pella). After 4 min incubation, grids were rinsed twice with 5 μ l distilled water and stained with 2% uranyl acetate for 1 min. Specimens were examined on an FEI T12 electron microscope.

Fibril diffraction. Fibrils formed from purified α -synuclein with and without N-terminal acetylation were concentrated by centrifugation, washed, and oriented while drying between two glass capillaries. Likewise, NACore nanocrystals were also concentrated, washed with nanopure water, and allowed to orient while drying between two glass capillaries. The glass capillaries holding the aligned fibrils or nanocrystals were mounted on a brass pin for diffraction at room temperature using 1.54 Å X-rays produced by a Rigaku FRE+ rotating anode generator equipped with an HTC imaging plate. All patterns were collected at a distance of 180 mm and analysed using the Adxv software package⁵⁷. A simulated pattern from the full length α -synuclein model presented in Fig. 3 was obtained by calculating structure factors from the model using the sfall module from CCP4, assigning the model a unit cell of $200 \times 4.74 \times 200$ Å. Cylindrical averaging of these structure factors about the fibre axis (y axis) direction produced a set of simulated fibril diffraction intensities.

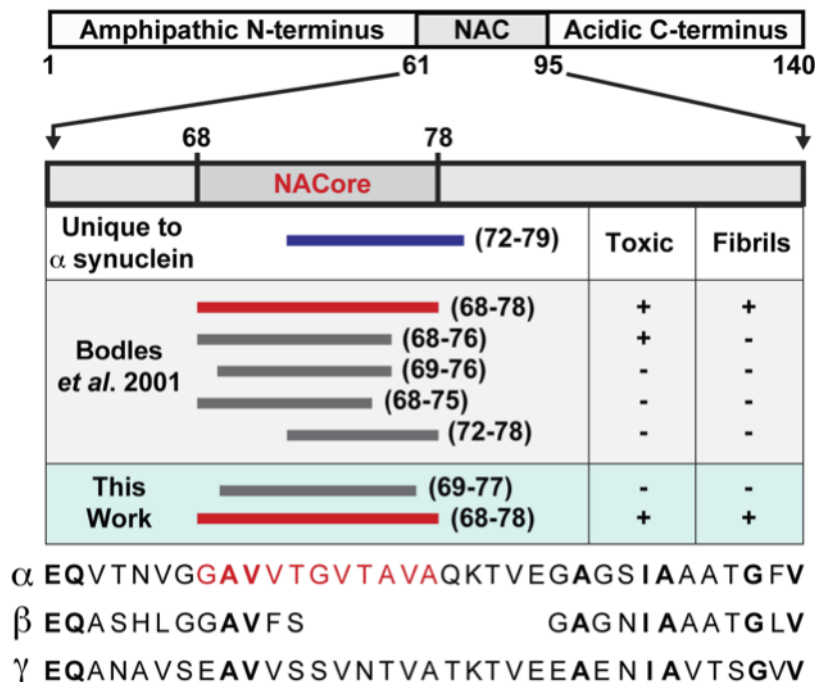
Cytotoxicity assays. Adherent PC12 cells (ATCC CRL-1721) were cultured in ATCC-formulated RPMI-1640 medium (ATCC 30-2001) supplemented with 10% horse serum and 5% fetal bovine serum and plated at 10,000 cells per well to a final volume of 90 μ l. All MTT assays were performed with Cell Titer 96 aqueous non-radioactive cell proliferation kit (MTT, Promega cat. no. 4100). Cells were cultured in 96-well plates for 20 h at 37 °C in 5% CO₂ before addition of samples (Costar cat. no. 3596). 10 μ l of sample was added to each well containing 90 μ l of medium and incubated for 24 h at 37 °C in 5% CO₂. Then, 15 μ l dye solution (Promega cat. no. 4102) was added into each well, followed by incubation for 4 h at 37 °C in 5% CO₂. This was followed by the addition of 100 μ l solubilization Solution/Stop Mix (Promega cat. no. 4101) to each well. After 12 h incubation at room temperature, the absorbance was measured at 570 nm. Background absorbance was recorded at 700 nm. The data was normalized with cells treated with 1% (w/v) SDS to 0% reduction, and cells treated with sample buffer to 100% reduction.

Lactose dehydrogenase assays were done using CytoTox-ONE Homogeneous Membrane Integrity, (Promega, cat. no. G7890) as per manufacturer's instructions.

In brief, cells were plated in 96-well, black-wall, clear-bottom (Fisher cat. no. 07-200-588) tissue culture plates at 10,000 cells per well to a final volume of 90 μ l. Cells were incubated for an additional 20 h at 37 °C in 5% CO₂ before addition of samples. Next, 10 μ l of sample was added to each well, following which the cells were incubated for another 24 h. 100 μ l of reagent was added to each well and incubated for 15 min at room temperature. The addition of 50 μ l of stop solution stopped the reaction. Fluorescence was measured in a Spectramax M5 (Molecular Devices) using excitation and emission wavelengths of 560 nm and 590 nm, respectively. Data was normalized using cells treated with buffer as 0% release and 0.1% Triton X-100 as 100% release.

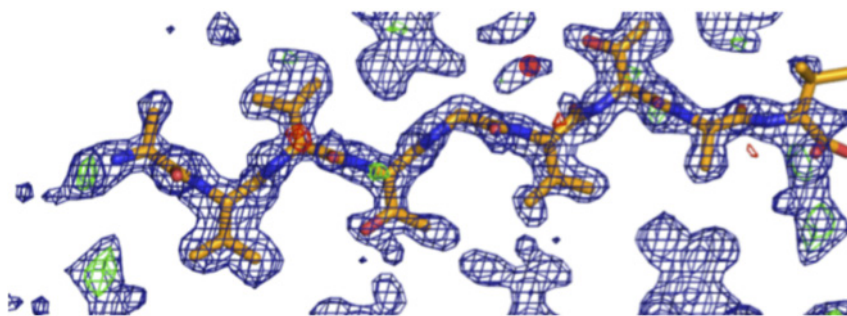
Construction of α -synuclein A53T fibril model. A model for full-length α -synuclein A53T mutant fibrils that are involved in the early onset of PD was constructed using a section of the NACore crystal packing as a scaffold. Figure 3 illustrates the four copies of the NACore segment used for the scaffold. The crystal structure of the two inner strands was adapted with minimal changes as the analogous segments 68–78. The structure of PreNAC was matched onto the weak interface of the NACore structure. Only 4 of the 11 side chains in the segment 46–56 differ from those in the NACore segment 68–78 and residues V51–V55 can be closely matched to V71–V75. Hence the model for both the homotypic interface and heterotypic interface in the full-length fibre model closely resemble those observed in the NACore structure. The regions outside these segments were adapted from the structure of the native α -synuclein fold, PDB ID: 2KKW⁵⁸. These segments were spliced in manually using COOT. The models were energy minimized and temperature annealed using the program CNS⁵⁹ with hydrogen-bonding potential⁶⁰. The simulated fibre diffraction pattern calculated from this model shows prominent reflections that agree with those observed in fibre diffraction patterns of NACore, α -synuclein, and N-acetyl α -synuclein (Extended Data Table 2).

43. Otwinowski, Z. & Minor, W. in *Methods in Enzymology* (eds Carter, C. W. Jr & Sweet, R. M.) Vol. 276 307–326 (Academic Press, 1997).
44. Weierstall, U., Spence, J. C. H. & Doak, R. B. Injector for scattering measurements on fully solvated biospecies. *Rev. Sci. Instrum.* **83**, 035108 (2012).
45. Hattne, J. *et al.* Accurate macromolecular structures using minimal measurements from X-ray free-electron lasers. *Nature Methods* **11**, 545–548 (2014).
46. Sauter, N. K., Hattne, J., Grosse-Kunstleve, R. W. & Echols, N. New Python-based methods for data processing. *Acta Crystallogr. D* **69**, 1274–1282 (2013).
47. Kabsch, W. XDS. *Acta Crystallogr. D* **66**, 125–132 (2010).
48. McCoy, A. J. *et al.* Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674 (2007).
49. Murshudov, G. N., Vagin, A. A. & Dodson, E. J. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D* **53**, 240–255 (1997).
50. Afonine, P. V. *et al.* Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr. D* **68**, 352–367 (2012).
51. Blanc, E. *et al.* Refinement of severely incomplete structures with maximum likelihood in BUSTER-TNT. *Acta Crystallogr. D* **60**, 2210–2221 (2004).
52. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D* **66**, 486–501 (2010).
53. Delano, W. *The PyMOL Molecular Graphics System* (Schrödinger LLC) <http://www.pymol.org>.
54. Jakes, R., Spillantini, M. G. & Goedert, M. Identification of two distinct synucleins from human brain. *FEBS Lett.* **345**, 27–32 (1994).
55. Johnson, M., Coulton, A. T., Gees, A. & Mulvihill, D. P. Targeted amino-terminal acetylation of recombinant proteins in *E. coli*. *PLoS ONE* **5**, e15801 (2010).
56. Whitelegge, J. P., Zhang, H., Aguilera, R., Taylor, R. M. & Cramer, W. A. Full subunit coverage liquid chromatography electrospray ionization mass spectrometry (LCMS+) of an oligomeric membrane protein: cytochrome *b₆f* complex from spinach and the cyanobacterium *Mastigocladus laminosus*. *Mol. Cell. Proteomics MCP* **1**, 816–827 (2002).
57. Arvai, A. *Adxv - A Program to Display X-ray Diffraction Images* (2015).
58. Rao, J. N., Jao, C. C., Hegde, B. G., Langen, R. & Ulmer, T. S. A combinatorial NMR and EPR approach for evaluating the structural ensemble of partially folded proteins. *J. Am. Chem. Soc.* **132**, 8657–8668 (2010).
59. Brunger, A. T. Version 1.2 of the crystallography and NMR system. *Nature Protocols* **2**, 2728–2733 (2007).
60. Fabiola, F., Bertram, R., Korostelev, A. & Chapman, M. S. An improved hydrogen bond potential: impact on medium resolution protein structures. *Protein Sci.* **11**, 1415–1423 (2002).
61. Read, R. J. Improved Fourier coefficients for maps using phases from partial structures with errors. *Acta Crystallogr. A* **42**, 140–149 (1986).



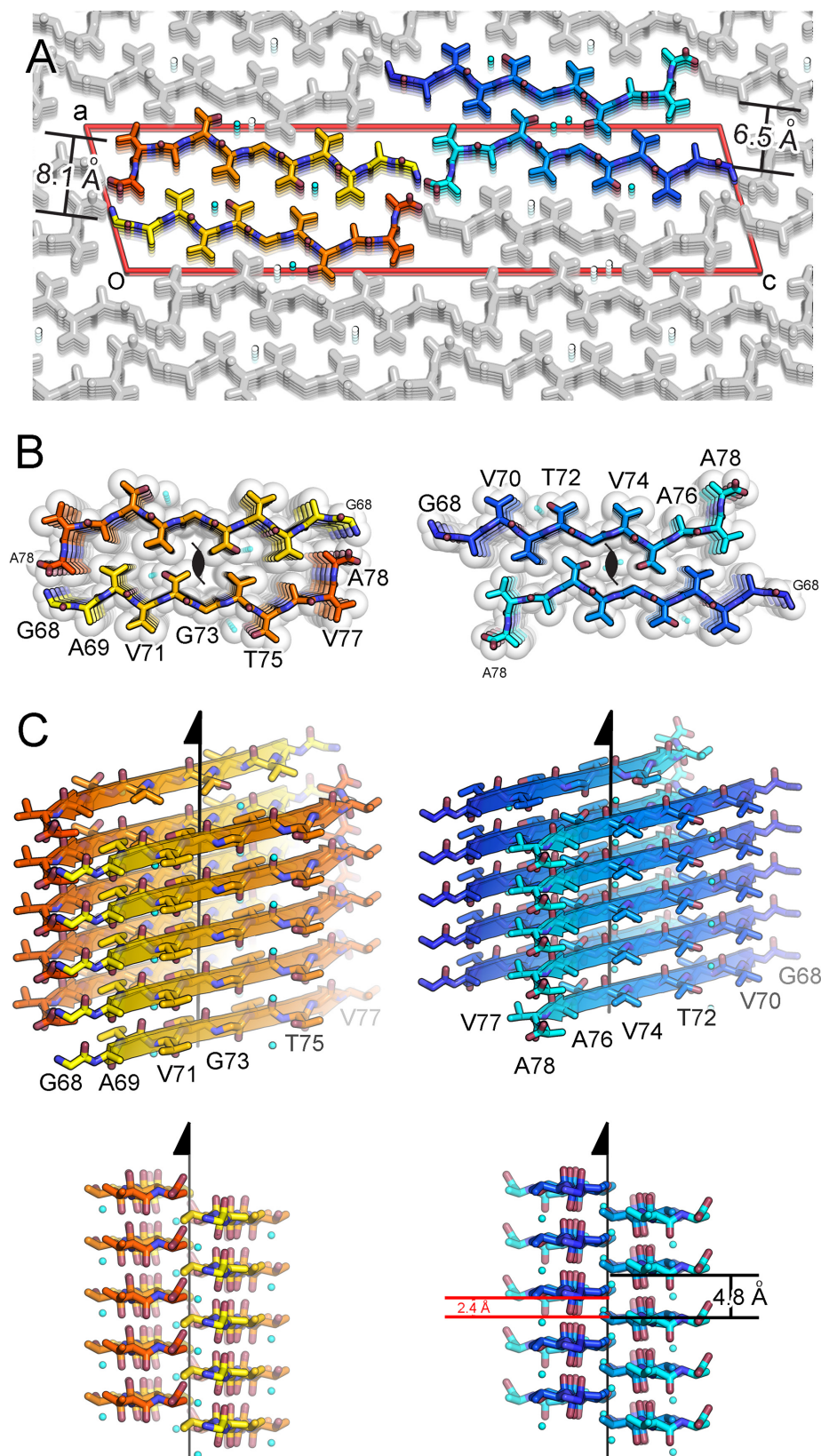
Extended Data Figure 1 | A schematic representation of α -synuclein, highlighting the NAC region (residues 61–95) and within it the NACore sequence (residues 68–78). A series of bars span regions of α -synuclein that are of interest to this work. Among the three synuclein paralogues (α , β and γ), the region whose sequence is unique to α -synuclein is shown as a blue bar (residues 72–83) that overlaps with a large portion of NACore. Segments investigated in ref. 23 are also shown. These span a variety of regions within

NACore. Two of the segments we investigate here, SubNACore and NACore, are shown in this context. Only one of the segments studied ref. 23 is an exact match to our NACore sequence, and only this segment is both toxic and fibrillar. The sequences of α -synuclein, β -synuclein, and γ -synuclein are shown as a reference with conserved residues in bold and the NACore sequence in red.



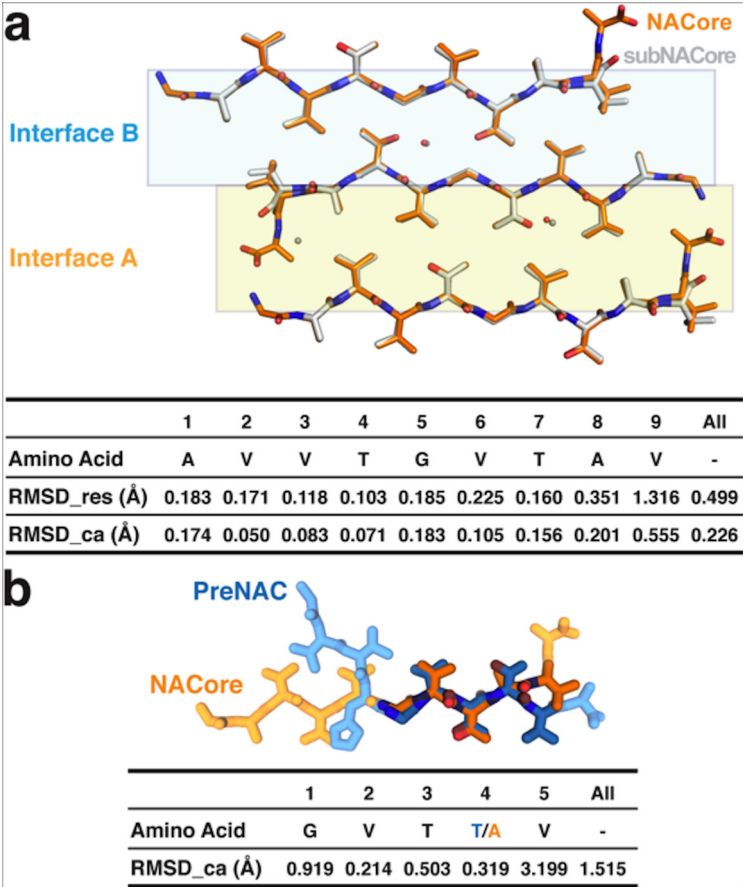
Extended Data Figure 2 | NACore difference density maps calculated after successful molecular replacement using the SubNACore search model clearly revealed the positions of the missing residues (positive $F_o - F_c$ density at N and C termini corresponding to G68 and A78) and one water

molecule near a threonine side chain (red circle); a second water was located during the refinement process. The blue mesh represents $2F_o - F_c$ density contoured at 1.2σ . The green and red mesh represent $F_o - F_c$ densities contoured at 3.0 and -3.0σ , respectively. All maps were σ_A -weighted⁶¹.



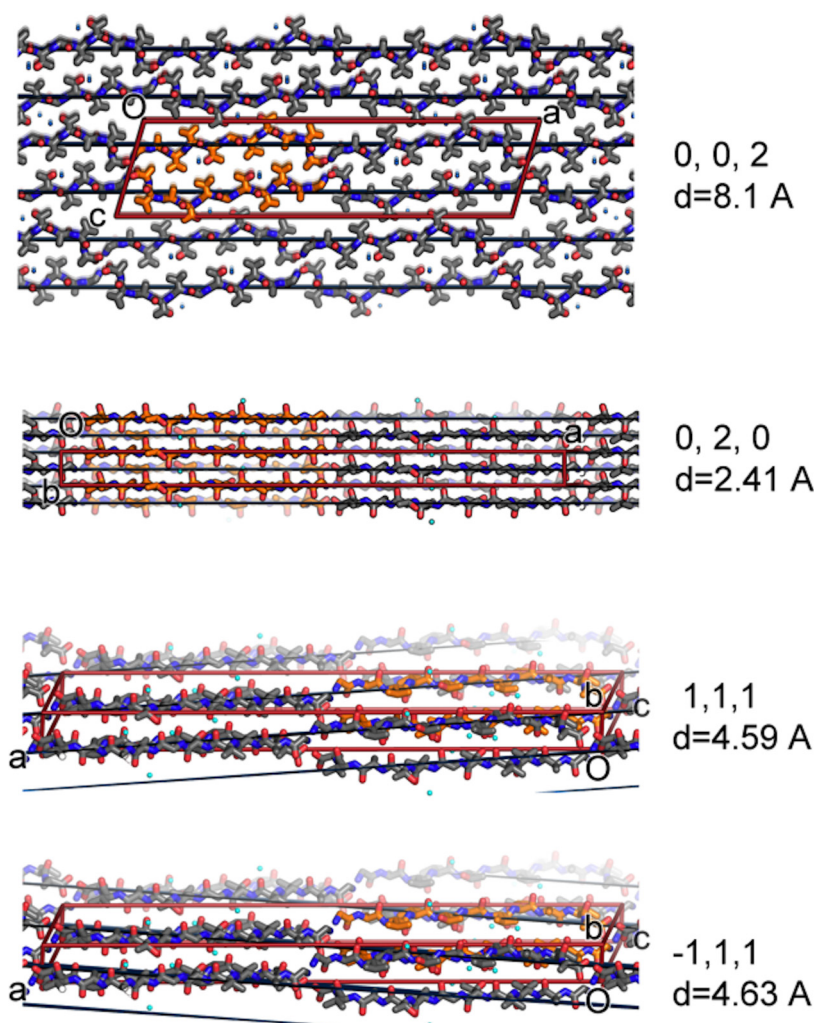
Extended Data Figure 3 | The crystal structure of NACore reveals pairs of sheets as in the spines of amyloid fibrils. **a**, NACore's two types of sheet-sheet interfaces: a larger interface (orange, 268 \AA^2 of buried accessible surface area per chain) we call interface A, and a weaker interface (blue, 167 \AA^2) we call interface B. The crystal is viewed along the hydrogen-bonding direction (crystal 'b' dimension). The red lines outline the unit cell. **b**, The van der Waals packing between sheets. The sheets are related by a 2_1 screw axis denoted in

black. The only gaps left by the interface are filled with water molecules which hydrogen-bond to the threonine residues (partially showing aqua spheres). The shape complementarity of both interfaces is 0.7. The viewing direction is the same as in **a**. **c**, Orthogonal view of the fibrillar assembly. The protofibril axis, coinciding with the 2_1 screw axis designated by the arrow, runs vertically between the pairs of sheets.



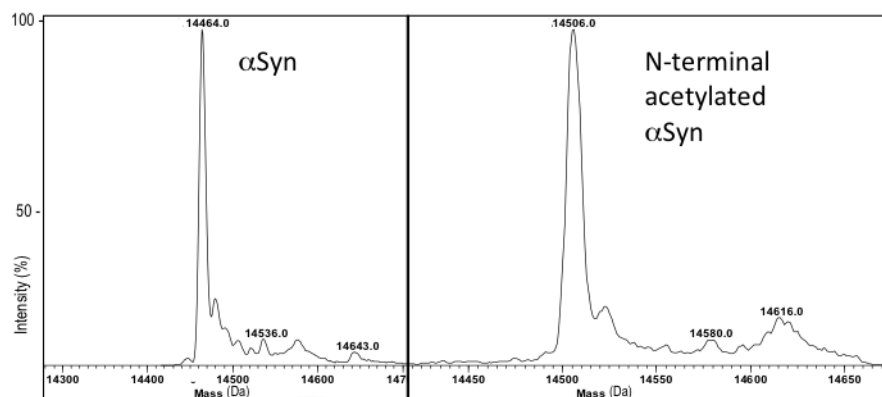
Extended Data Figure 4 | Comparison of the crystal packing for NACore and SubNACore. **a**, The face-to-face interactions are virtually the same for the pairs of NACore segments (orange chains) in its crystal structure and the SubNACore segments (white chains) in its structure (interfaces A and B shown in gold and blue, respectively). The table below shows the pairwise r.m.s.d. values comparing the nine residues shared in common between the structures.

RMSD_res is an all-atom comparison between residue pairs, while RMSD_ca compares only C α pairs. **b**, PreNAC (blue) is compared with NACore (orange). Five residues from each strand are shown in darker colour and the r.m.s.d. values between their C α pairs are compared in the table below. The PreNAC–NACore interaction mimics the weaker interface B in the NACore structure.



Extended Data Figure 5 | Intense reflections common among the NACore and the two polymorphs of full length α -synuclein suggest common structural features. Common structural features are illustrated here on the crystal packing diagrams of NACore. The (0,0,2) planes approximate the separation between sheets in interface A (orange). The (0,2,0), (-1,1,1), and (1,1,1) reflections are intense because the corresponding Bragg planes

reCAPITULATE the staggering of strands from opposing sheets. The red lines correspond to the unit cell boundaries and all planes are shown in black. The location of the unit cell origin is indicated by 'O'. The unit cell dimensions a, b, and c are labelled. Bragg spacings (spacings between planes), indicated by 'd', are indicated in angstroms.



Extended Data Figure 6 | Mass spectrometry analysis of recombinantly expressed, full-length α -synuclein, with and without N-terminal acetylation. The mass profile of wild-type full length α -synuclein (left) is compared to that of an N-terminally acetylated form of the protein (right). The mass shift for

the N-terminally acetylated form is appropriately shifted with respect to the native form of the protein (14464.0 Da for α -synuclein and 14506.0 Da for acetylated α -synuclein), within a margin of error of 4 Da.

Extended Data Table 1 | Statistics of data collection and atomic refinement for NACore, its fragment SubNACore, and PreNAC

Segment	SubNACore 69AVVTGVTAV ₇₇	NACore 68GAVVTGVTAVA ₅₆	PreNAC 47GVVHGVTVA ₅₆
Data collection			
Radiation source	Synchrotron	Electron	Electron
Space group	C2	C2	P21
Cell dimensions			
<i>a, b, c</i> (Å)	61.9, 4.80, 17.3	70.8, 4.82, 16.79	17.9, 4.7, 33.0
α, β, γ (°)	90, 104.1, 90	90, 105.7, 90	90, 94.3, 90
Resolution (Å)	1.85 (1.95-1.85)	1.43 (1.60-1.43)	1.41 (1.56-1.41)
Wavelength (Å)	0.9791	0.0251	0.0251
<i>R</i> _{merge}	0.117 (0.282)	0.173 (0.560)	0.236 (0.535)
<i>R</i> _{r.i.m.}	0.135 (0.322)	0.199 (0.647)	0.264 (0.609)
<i>R</i> _{p.i.m.}	0.065 (0.154)	0.093 (0.311)	0.185 (0.305)
<i>I</i> / σ <i>I</i>	5.2 (2.7)	5.5 (2.5)	4.6 (1.8)
CC _{1/2} (%)	99.5 (97.8)	99.4 (92.3)	96.7(74.0)
Completeness (%)	97.9 (98.3)	89.9 (82.6)	86.9 (69.6)
Multiplicity	4.1 (4.0)	4.4 (4.3)	3.7 (3.5)
Refinement			
Resolution (Å)	1.85 (2.07-1.85)	1.43 (1.60-1.43)	1.41 (1.41-1.57)
No. reflections	470 (125)	1073 (245)	1006 (239)
<i>R</i> _{work}	0.176 (0.248)	0.248 (0.253)	0.235 (0.336)
<i>R</i> _{free}	0.221 (0.286)	0.275 (0.331)	0.282 (0.329)
CC _{work}	0.964 (0.896)	0.947(0.618)	0.937(0.335)
CC _{free}	0.889 (0.993)	0.986(0.269)	0.967(0.361)
No. atoms			
Protein	57	66	66
Water	3	2	4
B-factors (Å ²)			
Protein	17.1	9.0	16.1
Water	27.6	2.7	24.6
Wilson B (Å ²)	11.8	10.3	13.8
R.m.s deviations			
Bond lengths (Å)	0.005	0.010	0.020
Bond angles (°)	1.1	1.6	2.0
PDB ID code	4RIK	4RIL	4ZNN
EMDB ID code		EMD-3028	EMD-3001

*Highest resolution shell is shown in parenthesis. Data quality is indicated by the redundancy independent merging *R* factor (r.i.m) and the precision indicating merging *R* factor (p.i.m.).

Extended Data Table 2 | Comparison of reflections observed in powder diffraction of fibrils of full-length α -synuclein, N-acetyl α -synuclein, and a synthetic pattern calculated from our α -synuclein model, to aligned nanocrystals of NACore.

Segment	Reflections (Å)
NACore GAVVTGVTAVA	2.21, 2.26, 2.39 , 2.52, 2.61, 2.68, 2.78, 3.02, 3.12, 3.34, 3.56, 3.86, 4.34, 4.57 , 5.16, 5.98, 7.56, 8.19 , 10.46, 11.63, 13.29, 16.61
α-syn	2.39 , 4.64 , 6.82, 8.29 , 10.06
N-acetyl α-syn	2.38 , 4.62 , 8.18 , 9.80, 11.90
Simulated α-syn	2.23, 2.25, 2.35 , 3.29, 3.63, 3.70, 3.95, 4.08, 4.56, 4.68 , 8.36 , 8.69, 21.76, 24.47, 27.61, 31.67

Bold reflections are strong and common to all three samples. Colours of the labelled reflections match those in Fig. 2.

Structure of the toxic core of α -synuclein from invisible crystals

Jose A. Rodriguez^{1*}, Magdalena I. Ivanova^{1*†}, Michael R. Sawaya^{1*}, Duilio Cascio^{1*}, Francis E. Reyes^{2*}, Dan Shi², Smriti Sangwan¹, Elizabeth L. Guenther¹, Lisa M. Johnson¹, Meng Zhang¹, Lin Jiang^{1†}, Mark A. Arbing¹, Brent L. Nannenga², Johan Hattne², Julian Whitelegge³, Aaron S. Brewster⁴, Marc Messerschmidt^{5†}, Sébastien Boutet⁵, Nicholas K. Sauter⁴, Tamir Gonen² & David S. Eisenberg¹

The protein α -synuclein is the main component of Lewy bodies, the neuron-associated aggregates seen in Parkinson disease and other neurodegenerative pathologies. An 11-residue segment, which we term NACore, appears to be responsible for amyloid formation and cytotoxicity of human α -synuclein. Here we describe crystals of NACore that have dimensions smaller than the wavelength of visible light and thus are invisible by optical microscopy. As the crystals are thousands of times too small for structure determination by synchrotron X-ray diffraction, we use micro-electron diffraction to determine the structure at atomic resolution. The 1.4 Å resolution structure demonstrates that this method can determine previously unknown protein structures and here yields, to our knowledge, the highest resolution achieved by any cryo-electron microscopy method to date. The structure exhibits protofibrils built of pairs of face-to-face β -sheets. X-ray fibre diffraction patterns show the similarity of NACore to toxic fibrils of full-length α -synuclein. The NACore structure, together with that of a second segment, inspires a model for most of the ordered portion of the toxic, full-length α -synuclein fibril, presenting opportunities for the design of inhibitors of α -synuclein fibrils.

The presynaptic protein α -synuclein, found in both soluble and membrane-associated fractions of the brain, aggregates in Parkinson disease (PD). These aggregates are the main component of Lewy bodies, the defining histological feature of this neurodegenerative disease, and have been shown to accompany neuronal damage¹. Two other observations point to aggregated α -synuclein as a molecular cause of PD². The first is that families with inherited forms of PD carry mutations in α -synuclein, such as A53T, and abundant Lewy bodies^{3–5}. The second is that families with duplicated or triplicated genes encoding α -synuclein develop early-onset PD, presumably because at high local concentrations α -synuclein is forced into amyloid^{6,7}.

Our focus is on a central segment of α -synuclein, residues 68–78, that we term NACore (Fig. 1), because of its critical role in both the aggregation and cytotoxicity of α -synuclein. NACore lies within a 35-residue domain of α -synuclein termed NAC (non-amyloid- β component, originally reported to be deposited with amyloid- β in the brains of Alzheimer's disease patients), which has been established as necessary and sufficient for the aggregation and toxicity of α -synuclein^{8–12} (Extended Data Fig. 1). For example, deletion of residues 71–82 prevents aggregation of α -synuclein *in vitro*, and abolishes both its aggregation and neurotoxicity in a *Drosophila* model of PD¹². Yet this segment in isolation from the rest of α -synuclein readily forms amyloid fibrils and is highly cytotoxic^{13,14}. Also, β -synuclein, the close homologue of α -synuclein, which does not aggregate and is not found in Lewy bodies, differs in sequence from α -synuclein principally by the lack of residues 74–84 that are part of NACore⁹.

Segments outside NAC also influence the aggregation of α -synuclein and have been associated with fibril structure^{15–17}. In brain extracts from

patients with multiple system atrophy, the core of α -synuclein fibrils extends approximately from residue 30 to 100¹⁸. Also the A53T mutation of α -synuclein can accelerate its transition into the amyloid state, and hence accelerate PD¹⁹. This mutation was found to induce the onset of PD at an early age²⁰, and consistent with this, α -synuclein containing this A53T mutation forms fibrils *in vitro* more rapidly than the wild type³. Thus we carried out screens for crystals of peptide segments within the NAC domain and adjacent regions, seeking structural information on the molecular basis of aggregation and toxicity of α -synuclein.

Structure determination by MicroED

Extensive crystal screens of two segments, NACore, residues 68GAVVTGVTAVA₇₈, and PreNAC, 47GVVHGVTVA₅₆, seemingly produced non-crystalline, amorphous aggregates. But upon examination using electron microscopy, we found the aggregates to be clusters of elongated nanocrystals only 50–300 nm in cross section and thus invisible by conventional light microscopy (Fig. 1). We confirmed well-ordered crystallinity of NACore at both the SACLA and LCLS free electron lasers. We also found that a nine-residue fragment within the NACore, which we term SubNACore, 69AVVTGVTAV₇₇, yielded crystals 1,000–10,000 times larger in volume than the NACore nanocrystals (Fig. 1). We were therefore able to apply synchrotron methods^{21,22} to these larger crystals to determine the structure of their amyloid-like fibrils. Although this nine-residue fragment is missing only two residues compared with NACore, it is not as toxic²³, offering some insight into the toxicity of α -synuclein, as described below.

To determine the structure of the invisible crystals of NACore and PreNAC, we turned to micro-electron diffraction (MicroED)^{24–26}. In

¹Howard Hughes Medical Institute, UCLA-DOE Institute, Departments of Biological Chemistry and Chemistry and Biochemistry, Box 951570, UCLA, Los Angeles, California 90095-1570, USA. ²Howard Hughes Medical Institute, Janelia Research Campus, 19700 Helix Drive, Ashburn, Virginia 20147, USA. ³Box 42, NPI-Semel Institute, 760 Westwood Plaza, UCLA, Los Angeles, California 90024, USA.

⁴Physical Biosciences Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720, USA. ⁵Linac Coherent Light Source, SLAC National Accelerator Laboratory, Menlo Park, California 94025, USA. [†]Present addresses: Department of Neurology and Program of Biophysics, University of Michigan School of Medicine, Ann Arbor, Michigan 48109, USA (M.I.); Department of Neurology, UCLA, Los Angeles, California 90095, USA (L.J.); National Science Foundation BioXFEL Science and Technology Center, Buffalo, New York 14203, USA (M.M.).

*These authors contributed equally to this work.

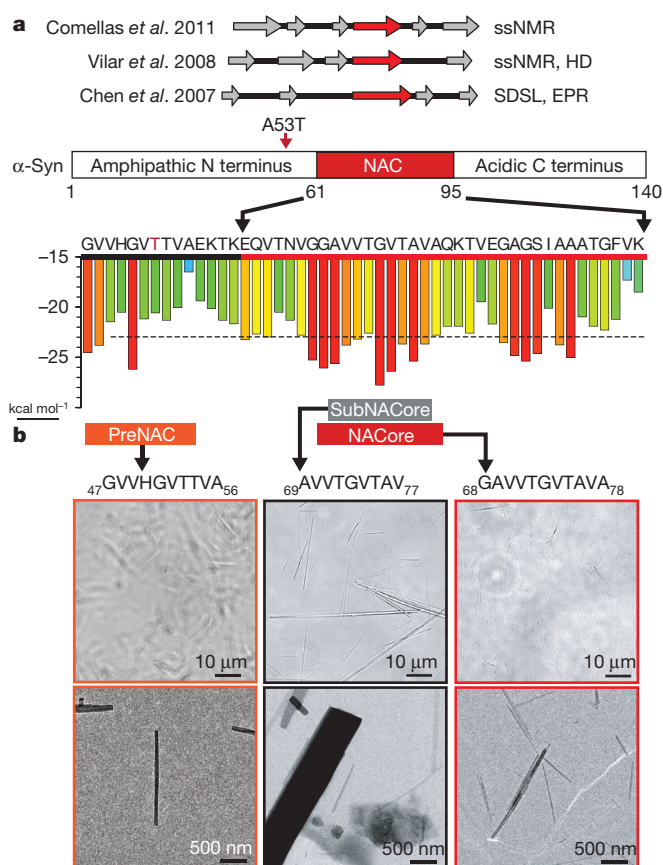


Figure 1 | NACore (residues 68–78) is the fibril-forming core of the NAC domain of full-length α -synuclein. **a**, Top, segments identified as β -strands by electron paramagnetic resonance (EPR), solid state nuclear magnetic resonance (ssNMR), hydrogen-deuterium exchange (HD), and site directed spin labelling (SDSL)^{17,40,41}. Bottom, predictions of the propensity of six-residue segments to form amyloid fibrils. The vertical axis indicates the propensity of steric zipper formation in Rosetta energy units⁴². Zipper-forming segments are predicted where bars cross the $-23 \text{ kcal mol}^{-1}$ threshold marked by the dashed line. Blue-to-red color gradient indicates weak-to-strong propensity of steric zipper formation. The A53T early-onset Parkinson mutation is indicated by a red arrow and red letter T. **b**, The miniscule size of the preNAC and NACore crystals used for MicroED is illustrated by this comparison to SubNACore microcrystals. Scale comparisons are illustrated on two magnifications using phase contrast light microscope images and electron micrographs, in which individual NACore and PreNAC nanocrystals are indistinguishable by light microscopy.

MicroED, an extremely low-dose electron beam is directed onto a nanocrystal within a transmission electron microscope under cryogenic conditions, yielding diffraction patterns such as those in Fig. 2. As the wavelength used in our experiments at 200 keV is very small (0.025 Å), the Ewald sphere is essentially flat, resulting in diffraction patterns that closely resemble a 2D slice through 3D reciprocal space. As the crystal is continuously rotated in the beam, a series of such diffraction patterns is collected²⁵. Scaling together diffraction data collected from multiple crystals produces a full 3D diffraction data set. MicroED has been successfully applied to the well-known structures of hen egg-white lysozyme^{25,26}, bovine liver catalase²⁷ and Ca^{2+} -ATPase²⁸. But NACore and PreNAC are the first previously unknown structures determined by MicroED.

For NACore and PreNAC, we collected MicroED patterns from nano-crystals that lay preferentially oriented, flat on the surface of a holey carbon Quantifoil grid, in a frozen-hydrated state. Grids were first screened for appropriately sized crystals, and candidate crystals screened for diffraction. We used crystals showing strong diffraction for data collection by continuous unidirectional rotation about a fixed

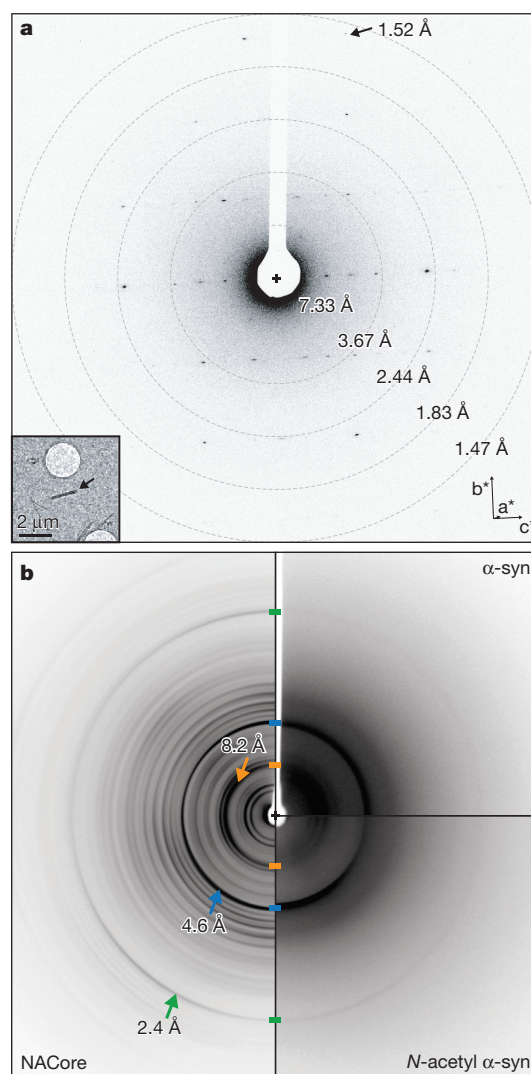


Figure 2 | Diffraction from NACore nanocrystals is similar to that from full length α -synuclein fibrils. **a**, Single-crystal electron diffraction pattern obtained during MicroED data collection (see text). Equally spaced concentric rings denote resolution shells. The highest resolution spot is at 1.52 Å (arrow). The inset shows the over-focused image of the diffracting crystal (arrow), which is $\sim 1,480 \times 200 \times 200 \text{ nm}$. Orientation of the reciprocal cell axes are indicated by the arrows labelled a^* , b^* , c^* . **b**, Composite of fibril diffraction patterns from α -synuclein (α -syn) preparations and NACore. Full-length α -synuclein reveals reflections that match those from NACore and N-terminally acetylated α -synuclein. The two patterns of full-length α -synuclein share with NACore three major peaks denoted by arrows: 8.2 Å (orange), 4.6 Å (blue), and 2.4 Å (green). The origin of these peaks can be traced to the (0,0,2), [(1,1,1), (-1,1,1)], and (0,2,0) planes in the NACore structure, respectively. We attribute the strong 8.2 Å reflection to the spacing between adjacent pairs of β -sheets.

axis, acquiring a series of diffraction frames at fixed time intervals²⁵. The needle-shaped crystals typically exceeded the length needed for MicroED; those that were unbent and 100 to 300 nm wide produced the best diffraction patterns. Data from multiple crystals were integrated, scaled and merged together (Extended Data Table 1).

The multi-crystal NACore and PreNAC data sets were phased by molecular replacement, using the atomic model of SubNACore and an ideal β -strand model, respectively, as probes. Residues of NACore which were missing from the SubNACore probe were clearly revealed in a difference density map calculated from NACore observed structure factor amplitudes and phases from the SubNACore probe structure (Extended Data Fig. 2). After subsequent refinement, two water

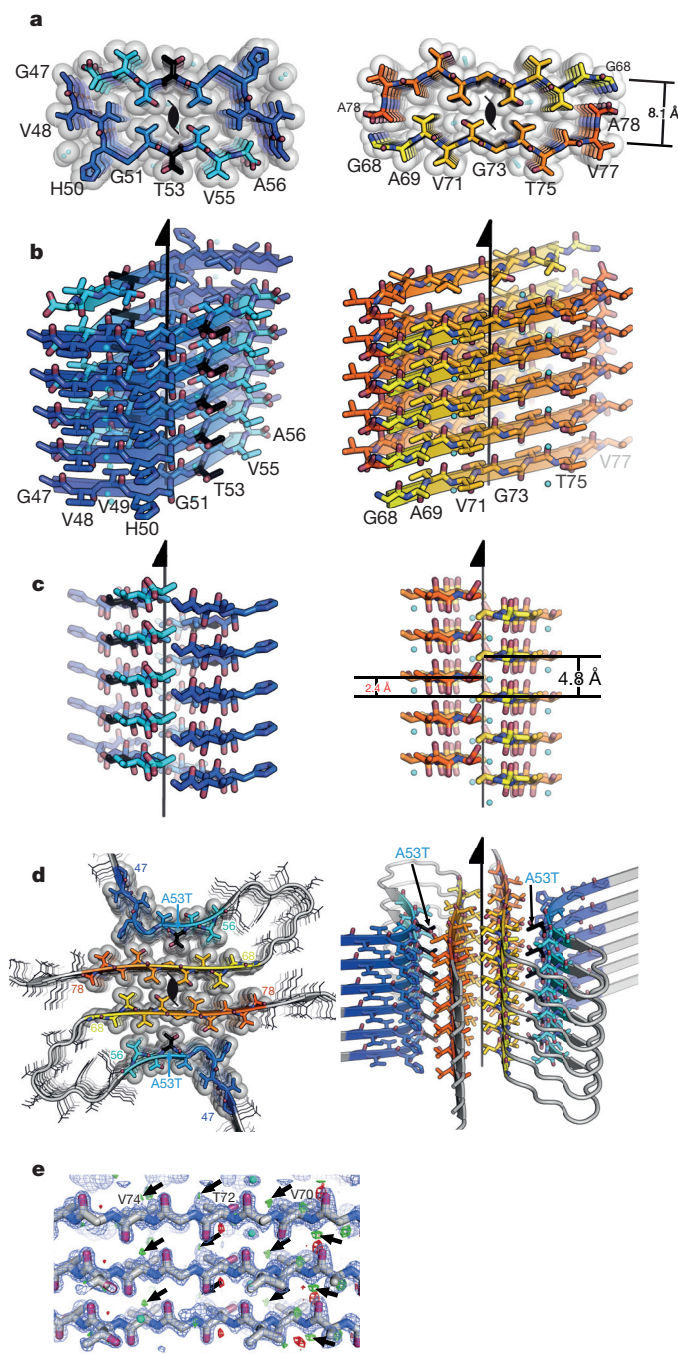


Figure 3 | Structure of the amyloid core of α -synuclein. **a**, The crystal structure of NACore (orange) reveals pairs of sheets as in the spines of amyloid fibrils. The A53T mutation (black) is shown in PreNAC (blue). The sheets in both structures are related by the 2_1 fibril axes shown in black. The gaps left by the interface are filled with water molecules which hydrogen-bond to the threonine residues (partially showing aqua spheres). **b**, **c**, Orthogonal views of the fibrillar assemblies. **d**, A speculative model of an α -synuclein protofibril containing the A53T mutation (black), where the strong interface of NACore (orange) forms the core of the fibril and its weaker interface interacts with PreNAC (blue). **e**, The locations of 5 out of a possible 73 protons are suggested by small, positive $F_o - F_c$ density (green contoured at 2.8σ , shown by arrows). The blue mesh is $2F_o - F_c$ density contoured at 1.4σ .

molecules, and several hydrogen atoms, were observed (Fig. 3e). Full models of NACore and PreNAC were refined against the MicroED data, producing structures at 1.4 \AA resolution with acceptable R factors (Extended Data Table 1). Electron scattering factors were used in the refinement calculations²⁹.

NACore structure

The structure of the NACore peptide chain is a nearly fully extended β -strand (Fig. 3 and Extended Data Fig. 3). These NACore strands stack in-register into β -sheets, as had been predicted by site-directed spin labelling^{16,17}. The sheets are paired (Fig. 3b), as is usual in amyloid spines, and the pairs of sheets form typical steric-zipper protofilaments (Fig. 3c), previously seen as the spines in many amyloid-like fibrils formed from short segments of fibril-forming proteins²¹. The unusual features of this steric zipper are that the 11-residue width of the zipper is longer than has been previously observed²², and each pair of sheets contains two water molecules, each associated with a threonine side chain within the interface. Most steric zippers are completely dry. Also, in our crystals of NACore, each sheet forms two snug interfaces: interface A, with 268 \AA^2 of buried accessible surface area per chain, is more extensive and presumably stronger than interface B (167 \AA^2), because the terminal residues of the chains in opposing sheets bend towards each other (Fig. 3 and Extended Data Fig. 4). The structure of PreNAC reveals a peptide chain that forms a β -strand kinked at Gly51. These strands are arranged into pairs of β -sheets that, like the NACore structure, interdigitate to form steric zipper protofilaments (Fig. 3). Of special note, a five-residue segment of PreNAC (₅₁GVTTV₅₅) differs in only one residue from a five-residue segment of NACore (₇₃GVTAV₇₇), and their α -carbons superimpose closely with a root mean square deviation (r.m.s.d.) of 1.5 \AA (Extended Data Fig. 4). This means that the weaker interface B of NACore mimics a hypothetical interface between NACore and PreNAC (Fig. 3d).

Relevance of NACore to Parkinson disease

The relationship of the structure of NACore to fibrils of full length α -synuclein is established by the resemblance of their diffraction patterns. Specifically, the fibre diffraction pattern of aligned fibrils of full-length and N-terminally acetylated³⁰ α -synuclein protein display the same principal peaks as the diffraction of aligned NACore nanocrystals (Fig. 2). All three fibrils display the strong reflection at 2.4 \AA in their diffraction patterns. As seen in Fig. 3 and Extended Data Fig. 5 this reflection arises in NACore because one β -sheet of the steric zipper is translated along the fibre axis with respect to the other β -sheet by 2.4 \AA , one half the 4.8 \AA spacing between β -strands, permitting the two sheets to interdigitate tightly together. All three share a strong 4.6 \AA reflection, which in NACore results from both the stacking of β -strands and the staggering between adjacent β -sheets of the steric zipper, while a shared reflection at near 8.2 \AA probably arises from the distance between the adjacent pairs of β -sheets that make up the α -synuclein fibril (Fig. 2 and Extended Data Fig. 5). This comparison of fibre diffraction patterns (Extended Data Table 2) strongly suggests that the structure of NACore is similar to the spine of the toxic fibrils of full α -synuclein.

The combined structures of NACore and PreNAC allow us to construct a speculative model for much of the ordered segments of the A53T early-onset mutant α -synuclein (Fig. 3d). Experimental support of this model comes from the agreement of its simulated fibre diffraction with the measured diffraction patterns of α -synuclein and N-acetyl α -synuclein fibrils, as well as aligned NACore nanocrystals (Extended Data Table 2). Above we hypothesized that the weaker interface B of NACore might mimic an intramolecular interaction of PreNAC with NACore (Fig. 3). In fact, the interacting side chains in the weaker NACore interface B (G73, T75 and V77) are identical to the side chains (G51, T53, V55) interacting in the hypothetical interface of PreNAC with NACore. Assuming that this interface actually forms in fibrils of the early-onset mutant A53T, we built the model shown in Fig. 3d. The hypothetical interface of this model offers a possible reason for a greater propensity of the A53T mutant to aggregate than the wild-type sequence, conceivably leading to the early onset of PD.

The identity and structure of the cytotoxic amyloid formed by α -synuclein remains a subject of intensive research^{19,31–35}. The weight

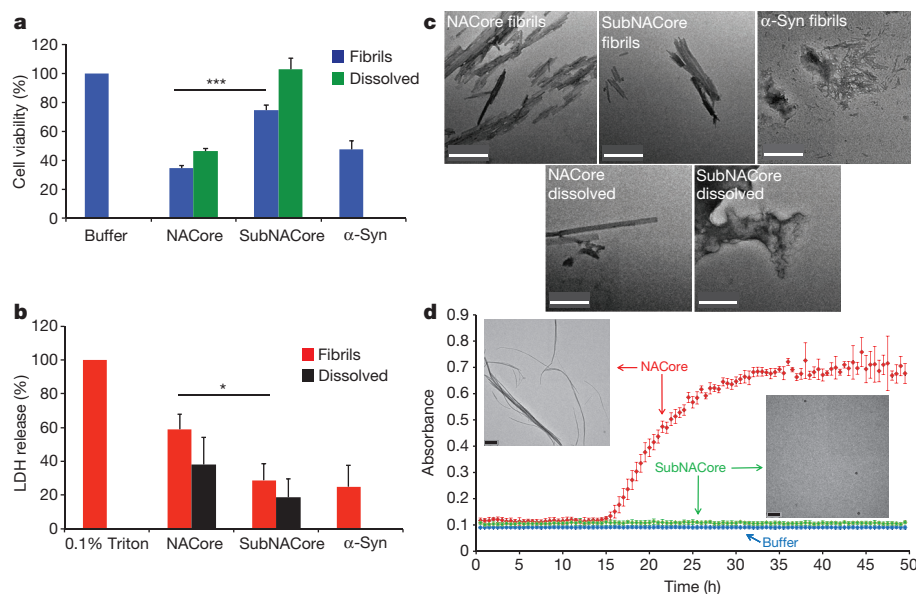


Figure 4 | NACore aggregates faster than SubNACore and is more cytotoxic to cultured cells. **a**, **b**, Cytotoxicity of NACore, SubNACore and α -synuclein measured on PC12 cells using a 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT) assay (**a**) and a lactate dehydrogenase (LDH)-release assay (**b**). In both assays NACore is more toxic than SubNACore. Also, shaken fibrils are more toxic than an equal concentration of freshly dissolved sample. Results shown as mean \pm s.e.m. based on triplicate samples. A *t*-test was used to measure statistical significance; **P* < 0.05, ****P* < 0.001. **c**, Representative electron micrographs of NACore, SubNACore

and α -synuclein (α -syn) samples tested for cytotoxicity. NACore and α -synuclein show abundant fibrils but SubNACore shows few. NACore also forms fibrils immediately upon dissolving, whereas SubNACore shows no fibres, but instead amorphous aggregates. Scale bar, 500 nm. **d**, NACore and SubNACore were aggregated in identical conditions and monitored by turbidity. NACore begins to aggregate after 15 h while SubNACore forms no aggregates for up to 50 h. Electron microscopy of the samples at 50 h confirmed the turbidity readings (insets; scale bars, 2 μ m), with error bars denoting standard deviation based on triplicate samples.

of evidence over the past decade has tilted scientific opinion from the fully developed amyloid fibrils found in Lewy bodies as the toxic entities to smaller, transient amyloid oligomers. Yet recently, quantitative arguments have been put forward in favour of fibrils³⁶. Our experiments of the cytotoxicity of NACore on PC12 cells (Fig. 4) are consistent with the view that fibrils are toxic: we find that NACore shaken and aggregated for 72 h displays abundant fibrils, is more toxic than freshly dissolved NACore (Fig. 4), and is comparably toxic to similarly aggregated full α -synuclein. We also find greater cytotoxicity of NACore than SubNACore, which is shorter by two residues. This is consistent with the more rapid fibril formation of NACore than of SubNACore (Fig. 4d). These observations do not rule out the formation of a non-fibrillar, oligomeric assembly, present, but undetected, in our aggregated samples of NACore and α -synuclein. Of course, NACore is merely a fragment of full length α -synuclein, and lacks most of the membrane-binding motifs of the N terminus of the protein, which have been implicated in membrane disruption^{37,38}. Yet it is clear that NACore is the minimum entity that recapitulates all the features of full length α -synuclein aggregation and toxicity.

MicroED diffraction of invisible crystals

The miniscule size of NACore crystals is typical of amyloid and also of various other biological crystals of interest. For amyloid crystals, our speculation is that the tiny size is a consequence of the natural twist of β -sheets that form the protofilaments of the fibrils. The crystal lattice restrains the twist, creating a strain in these crystals, which increases as crystals grow. Eventually this strain prevents further addition of β -strands, limiting the thickness of the needle crystals. In our experience, longer segments (for example, 11 residues compared to 9 residues) limit crystal growth even more; in the case of 11-residue NACore and 10-residue PreNAC, the strain produces nanocrystals, invisible by optical microscopy. These crystals are too small for mounting and conventional synchrotron data collection, but are ideally suited for analysis by MicroED. They are $\sim 10^{10}$ times smaller

than Perutz's haemoglobin crystals and $\sim 10^{12}$ times smaller than von Laue's CuSO_4 crystal, which yielded the first X-ray diffraction pattern. Our structures of NACore and PreNAC demonstrate that MicroED is capable of determining new and accurate structures of biological material at atomic resolutions. This finding paves the way for applications of MicroED to other biological substances of importance for which only nanocrystals can be grown. In our particular application, we have been able to learn the atomic arrangement of the core of the crucial NAC domain. This presents opportunities for structure-based design of inhibitors of amyloid formation of α -synuclein³⁹.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 30 October 2014; accepted 13 August 2015.

Published online 9 September 2015.

- Spillantini, M. G. *et al.* α -Synuclein in Lewy bodies. *Nature* **388**, 839–840 (1997).
- Goedert, M., Spillantini, M. G., Del Tredici, K. & Braak, H. 100 years of Lewy pathology. *Nature Rev. Neurol.* **9**, 13–24 (2013).
- Polymeropoulos, M. H. *et al.* Mutation in the α -synuclein gene identified in families with Parkinson's disease. *Science* **276**, 2045–2047 (1997).
- Krüger, R. *et al.* Ala30Pro mutation in the gene encoding α -synuclein in Parkinson's disease. *Nature Genet.* **18**, 106–108 (1998).
- Zarranz, J. J. *et al.* The new mutation, E46K, of α -synuclein causes Parkinson and Lewy body dementia. *Ann. Neurol.* **55**, 164–173 (2004).
- Ibáñez, P. *et al.* Causal relation between α -synuclein gene duplication and familial Parkinson's disease. *Lancet* **364**, 1169–1171 (2004).
- Singleton, A. B. *et al.* α -Synuclein locus triplication causes Parkinson's disease. *Science* **302**, 841 (2003).
- Ueda, K. *et al.* Molecular cloning of cDNA encoding an unrecognized component of amyloid in Alzheimer disease. *Proc. Natl Acad. Sci. USA* **90**, 11282–11286 (1993).
- Biere, A. L. *et al.* Parkinson's disease-associated α -synuclein is more fibrillogenic than β - and γ -synuclein and cannot cross-seed its homologs. *J. Biol. Chem.* **275**, 34574–34579 (2000).
- Giasson, B. I., Murray, I. V. J., Trojanowski, J. Q. & Lee, V. M.-Y. A hydrophobic stretch of 12 amino acid residues in the middle of α -synuclein is essential for filament assembly. *J. Biol. Chem.* **276**, 2380–2386 (2001).
- Du, H.-N. *et al.* A peptide motif consisting of glycine, alanine, and valine is required for the fibrillization and cytotoxicity of human α -synuclein. *Biochemistry* **42**, 8870–8878 (2003).

12. Periquet, M., Fulga, T., Myllykangas, L., Schlossmacher, M. G. & Feany, M. B. Aggregated α -synuclein mediates dopaminergic neurotoxicity *in vivo*. *J. Neurosci.* **27**, 3338–3346 (2007).
13. Han, H., Weinreb, P. H. & Lansbury, P. T. The core Alzheimer's peptide NAC forms amyloid fibrils, which seed and are seeded by β -amyloid: is NAC a common trigger or target in neurodegenerative disease? *Chem. Biol.* **2**, 163–169 (1995).
14. El-Agnaf, O. M. *et al.* Aggregates from mutant and wild-type α -synuclein proteins and NAC peptide induce apoptotic cell death in human neuroblastoma cells by formation of β -sheet and amyloid-like filaments. *FEBS Lett.* **440**, 71–75 (1998).
15. Crowther, R. A., Daniel, S. E. & Goedert, M. Characterisation of isolated α -synuclein filaments from substantia nigra of Parkinson's disease brain. *Neurosci. Lett.* **292**, 128–130 (2000).
16. Der-Sarkissian, A., Jao, C. C., Chen, J. & Langen, R. Structural organization of α -synuclein fibrils studied by site-directed spin labeling. *J. Biol. Chem.* **278**, 37530–37535 (2003).
17. Chen, M., Margittai, M., Chen, J. & Langen, R. Investigation of α -synuclein fibril structure by site-directed spin labeling. *J. Biol. Chem.* **282**, 24970–24979 (2007).
18. Miake, H., Mizusawa, H., Iwatsubo, T. & Hasegawa, M. Biochemical characterization of the core structure of α -synuclein filaments. *J. Biol. Chem.* **277**, 19213–19219 (2002).
19. Conway, K. A., Harper, J. D. & Lansbury, P. T. Accelerated *in vitro* fibril formation by a mutant α -synuclein linked to early-onset Parkinson disease. *Nature Med.* **4**, 1318–1320 (1998).
20. Polymeropoulos, M. H. *et al.* Mutation in the α -synuclein gene identified in families with Parkinson's disease. *Science* **276**, 2045–2047 (1997).
21. Nelson, R. *et al.* Structure of the cross- β spine of amyloid-like fibrils. *Nature* **435**, 773–778 (2005).
22. Sawaya, M. R. *et al.* Atomic structures of amyloid cross- β spines reveal varied steric zippers. *Nature* **447**, 453–457 (2007).
23. Bodles, A. M., Guthrie, D. J., Greer, B. & Irvine, G. B. Identification of the region of non-A β component (NAC) of Alzheimer's disease amyloid responsible for its aggregation and toxicity. *J. Neurochem.* **78**, 384–395 (2001).
24. Nannenga, B. L. & Gonen, T. Protein structure determination by MicroED. *Curr. Opin. Struct. Biol.* **27**, 24–31 (2014).
25. Nannenga, B. L., Shi, D., Leslie, A. G. W. & Gonen, T. High-resolution structure determination by continuous-rotation data collection in MicroED. *Nature Methods* **11**, 927–930 (2014).
26. Shi, D., Nannenga, B. L., Iadanza, M. G. & Gonen, T. Three-dimensional electron crystallography of protein microcrystals. *eLife* **2**, e01345 (2013).
27. Nannenga, B. L., Shi, D., Hattne, J., Reyes, F. E. & Gonen, T. Structure of catalase determined by MicroED. *eLife* **3**, e03600 (2014).
28. Yonekura, K., Kato, K., Ogasawara, M., Tomita, M. & Toyoshima, C. Electron crystallography of ultrathin 3D protein crystals: atomic model with charges. *Proc. Natl Acad. Sci. USA* **112**, 3368–3373 (2015).
29. Doyle, P. A. & Turner, P. S. Relativistic Hartree–Fock X-ray and electron scattering factors. *Acta Crystallogr. A* **24**, 390–397 (1968).
30. Sarafian, T. A. *et al.* Impairment of mitochondria in adult mouse brain overexpressing predominantly full-length, N-terminally acetylated human α -synuclein. *PLoS ONE* **8**, e63557 (2013).
31. Caughey, B. & Lansbury, P. T. Protofibrils, pores, fibrils, and neurodegeneration: separating the responsible protein aggregates from the innocent bystanders. *Annu. Rev. Neurosci.* **26**, 267–298 (2003).
32. Danzer, K. M., Schnack, C., Sutcliffe, A., Hengerer, B. & Gillardon, F. Functional protein kinase arrays reveal inhibition of p-21-activated kinase 4 by α -synuclein oligomers. *J. Neurochem.* **103**, 2401–2407 (2007).
33. Karpinar, D. P. *et al.* Pre-fibrillar α -synuclein variants with impaired β -structure increase neurotoxicity in Parkinson's disease models. *EMBO J.* **28**, 3256–3268 (2009).
34. Winner, B. *et al.* *In vivo* demonstration that α -synuclein oligomers are toxic. *Proc. Natl Acad. Sci. USA* **108**, 4194–4199 (2011).
35. Chen, S. W. *et al.* Structural characterization of toxic oligomers that are kinetically trapped during α -synuclein fibril formation. *Proc. Natl Acad. Sci. USA* **112**, E1994–E2003 (2015).
36. Bousset, L. *et al.* Structural and functional characterization of two α -synuclein strains. *Nature Commun.* **4**, 2575 (2013).
37. Auluck, P. K., Caraveo, G. & Lindquist, S. α -Synuclein: membrane interactions and toxicity in Parkinson's disease. *Annu. Rev. Cell Dev. Biol.* **26**, 211–233 (2010).
38. Lee, J. C., Langen, R., Hummel, P. A., Gray, H. B. & Winkler, J. R. α -Synuclein structures from fluorescence energy-transfer kinetics: implications for the role of the protein in Parkinson's disease. *Proc. Natl Acad. Sci. USA* **101**, 16466–16471 (2004).
39. Sievers, S. A. *et al.* Structure-based design of non-natural amino-acid inhibitors of amyloid fibril formation. *Nature* **475**, 96–100 (2011).
40. Comellas, G. *et al.* Structured regions of α -synuclein fibrils include the early-onset Parkinson's disease mutation sites. *J. Mol. Biol.* **411**, 881–895 (2011).
41. Vilar, M. *et al.* The fold of α -synuclein fibrils. *Proc. Natl Acad. Sci. USA* **105**, 8637–8642 (2008).
42. Goldschmidt, L., Teng, P. K., Riek, R. & Eisenberg, D. Identifying the amyloids, proteins capable of forming amyloid-like fibrils. *Proc. Natl Acad. Sci. USA* **107**, 3487–3492 (2010).

Acknowledgements We thank C. Liu for supplying PC12 cells; APS staff for beam line help solving SubNACore: M. Capel, K. Rajashankar, N. Sukumar, J. Schuermann, I. Kourinov and F. Murphy at NECAT beam lines 24-ID at APS funded by the National Institute of General Medical Sciences from the National Institutes of Health (P41 GM103403) and the DOE Office of Science by Argonne National Laboratory under Contract No. DE-AC02-06CH11357. We thank the LCLS injection staff support: S. Botha, R. Shoeman and I. Schlichting. A.S.B. and N.K.S. were supported by NIH grants GM095887 and GM102520 and by the Director, Office of Science, Department of Energy (DOE) under contract DE-AC02-05CH11231 for data-processing methods. This work was supported by the US Department of Energy Office of Science, Office of Biological and Environmental Research program under award number DE-FC02-02ER63421. We also acknowledge the award MCB-0958111 from the National Science Foundation, award 1R01-AG029430 from the National Institutes of Health, award NIH-AG016570 from Alzheimer's Disease Research (ADRC) at UCLA, and HHMI for support. J.A.R. was supported by the Giannini Foundation.

Author Contributions M.I.I. characterized the α -synuclein segments and crystals. M.I.I. and S.S. conducted the toxicity assays. L.M.J. synthesized and purified NACore peptide. M.A.A. prepared the N-terminally acetylated α -synuclein. S.S. and M.Z. prepared wild-type α -synuclein. J.W. performed the mass spectrometry analyses of α -synuclein. M.I.I. and L.M.J. crystallized NACore. E.G. grew crystals of SubNACore. E.G., M.I.I. and M.R.S. collected and processed the data and solved the structure of SubNACore. L.J. and J.A.R. identified and crystallized PreNAC. J.A.R., D.S., B.L.N. and T.G. collected MicroED data on PreNAC and NACore nanocrystals. J.A.R., F.E.R., J.H., T.G., L.J., M.R.S., and D.C. processed the MicroED data and solved the structure of PreNAC and NACore. J.A.R., M.R.S., D.C., M.M. and S.B. collected XFEL diffraction from NACore nanocrystals. A.S.B. and N.K.S. processed the XFEL data. M.R.S. and L.J. built the structure model of A53T α -synuclein protofibril. J.A.R., M.I.I., M.R.S., D.C., S.S. and E.G. prepared the figures. J.A.R., M.I.I., M.R.S., D.C., T.G. and D.S.E. wrote the paper, and all authors commented on the paper.

Author Information Atomic coordinates and structure factors have been deposited in the Protein Data Bank under accession codes 4RIK (SubNACore), 4RIL (NACore) and 4ZNN (PreNAC). The maps for PreNAC and NACore have been deposited in the EMDB with accession codes EMD-3001 and EMD-3028, respectively. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to D.S.E. (david@mbi.ucla.edu) and T.G. (gonent@janelia.hhmi.org).

METHODS

Data reporting. No statistical methods were used to predetermine sample size. The experiments were not randomized. The investigators were not blinded to allocation during experiments and outcome assessment.

Crystallization. Microcrystals of SubNACore ($_{69}$ AVVTGVTAV $_{77}$) were grown from synthetic peptide purchased from CS Bio. Crystals were grown at room temperature by hanging drop vaporization. Lyophilized peptide was dissolved in water at 2.9 mg ml^{-1} concentration in 48 mM lithium hydroxide. Peptide was mixed in a 2:1 ratio with reservoir containing 0.9 M ammonium phosphate, and 0.1 M sodium acetate pH 4.6.

Nanocrystals of NACore, $_{68}$ GAVVTGVTAV $_{78}$, were grown from synthetic peptide purchased from CS Bio. Ten batches of synthesized peptide (CSBio) at a concentration of 1 mg ml^{-1} in sterile water were shaken at 37°C on a Torrey Pines orbital mixing plate at speed setting 9, overnight. The insoluble material was washed in 30% (w/v) glycerol then stored in water at room temperature before diffraction. The sample contained a mixture of fibrils and crystals.

Nanocrystals of PreNAC ($_{47}$ GVVHGVTTVA $_{56}$) were grown from synthetic peptide purchased from InnoPep. Crystallization trials of synthesized peptide were prepared in batch. Peptide was weighed and dissolved in sterile-filtered 50 mM phosphate buffer pH 7.0 with 0.1% DMSO at a concentration of 5 mg ml^{-1} . This solution was shaken at 37°C on a Torrey Pines orbital mixing plate at speed setting 9, overnight.

Data collection and processing. X-ray diffraction data from microcrystals of SubNACore were collected using synchrotron radiation at the Advanced Photon Source, Northeast Collaborative Access Team micro focus beam line 24-ID-E. The beam line was equipped with an ADSC Quantum 315 CCD detector. Data from a single crystal were collected in 5° wedges at a wavelength of 0.9791 \AA using a $5 \mu\text{m}$ beam diameter. We used data from three different sections along the needle axis. The crystals were cryo-cooled (100 K) for data collection. Data were processed and reduced using Denzo/Scalepack from the HKL suite of programs⁴³.

Electron diffraction data from nanocrystals of NACore and PreNAC were collected using MicroED techniques^{25,26}. These nanocrystals typically clump together. To break up the clumps, an approximately $100 \mu\text{l}$ volume of nanocrystals was placed in a sonication bath for 30 min. Nanocrystals were deposited onto a Quantifoil holey-carbon EM grid in a $2\text{--}3 \mu\text{l}$ drop after appropriate dilution, which optimized for crystal density on the grid. All grids were then blotted and vitrified by plunging into liquid ethane using a Vitrobot Mark IV (FEI), then transferring to liquid nitrogen for storage. Frozen hydrated grids were transferred to a cryo-TEM using a Gatan 626 cryo-holder. Diffraction patterns and crystal images were collected using an FEG-equipped FEI Tecnai F20 TEM operating at 200 kV and recorded using a bottom mount TVIPS F416 CMOS camera with a sensor size of $4,096 \times 4,096$ pixels, each $15.6 \times 15.6 \mu\text{m}$. Diffraction patterns were recorded by operating the detector in rolling shutter mode with 2×2 pixel binning, producing a final image $2,048 \times 2,048$ pixels in size. Individual image frames were taken with exposure times of $3\text{--}4 \text{ s}$ per image, using a selected area aperture with an illuminating spot size of approximately $1 \mu\text{m}$. This geometry equates to an electron dose of less than $0.1 \text{ e}^- \text{ per \AA}^2 \text{ per second}$. During each exposure, crystals were continuously rotated within the beam at a rate of 0.3° per second, corresponding to 1.2° wedge per frame. Diffraction data were collected from several crystals each oriented differently with respect to the rotation axis. These data sets each spanned wedges of reciprocal space ranging from 40° to 80° .

X-ray diffraction data from nanocrystals of NACore were collected using XFEL radiation at the CXI instrument (Coherent X-ray Imaging) at the Linear Coherent Light Source (LCLS)-SLAC. The photon energy of the X-ray pulses was 8.52 keV (1.45 \AA). Each 40 fs pulse contained up to 6×10^{11} photons at the sample position, taking into account a beam line transmission of 60%. The diameter of the beam was approximately $1 \mu\text{m}$. We used a concentration of approximately $25 \mu\text{l}$ of pelleted material suspended in 1 ml water. The sample was injected into the XFEL beam using a liquid jet injector and a gas dynamic virtual nozzle⁴⁴. The micro jet width was approximately $4 \mu\text{m}$ and the flow rate was $40 \mu\text{l min}^{-1}$. The sample caused noticeable sputtering of the liquid jet. XFEL data were processed using cctbx.xfel^{45,46}.

Calibration of the sample to detector distance in MicroED was accomplished using a polycrystalline gold standard and by referencing the prominent reflections in the electron diffraction experiment with the corresponding reflections in the XFEL data. Calibration of the x/y locations of the 64-tile CSPAD detector was performed by cctbx.xfel by refining the optically measured tile positions against a thermolysin data set⁴⁵.

To gain compatibility with conventional X-ray data processing programs, the MicroED diffraction images were converted from tiff or TVIPS format to the SMV crystallographic format. We used XDS to index the diffraction images⁴⁷, and XSCALE for merging and scaling together data sets originating from different

crystals. For NACore, data from four crystals were merged, while for PreNAC, data from three crystals were merged to assemble the final data sets (see Extended Data Table 1).

Structure determination. The molecular replacement solution for SubNACore was obtained using the program Phaser⁴⁸. The search model consisted of a geometrically ideal β -strand composed of nine alanine residues. Crystallographic refinements were performed with the program Refmac⁴⁹.

The molecular replacement solution for NACore was obtained using the program Phaser⁴⁸. The search model consisted of the SubNACore structure determined previously. Crystallographic refinements were performed with the program Phenix⁵⁰ and Buster⁵¹.

The molecular replacement solution for PreNAC was obtained using the program Phaser⁴⁸. The search model consisted of a geometrically ideal β -strand composed of six residues with sequence GVTVA. Crystallographic refinements were performed with the program Phenix⁵⁰ and Refmac⁴⁹.

Model building for all segments was performed using COOT⁵². Data processing and refinement statistics are reported in Extended Data Table 1. The coordinates of the final models and the structure factors have been deposited in the Protein Data Bank with PDB code 4RIK for SubNACore, 4RIL for NACore, and 4ZNN for PreNAC. The structures were illustrated using Pymol⁵³.

Protein expression and purification. The human wild-type α -synuclein construct has been previously characterized⁵⁴ (pRK172, ampicillin, T7 promoter) with sequence: MDVFMKGLSKAKEGVVAAAEKTKQGVAAEAGKTKGVLYVGSKTKEGVVHGVATVAEKTKEQVTNVGGAVVTGVTAVAQKTEVGAGSIAAATGFVKDKQLGKNEEGAPQEGILEMPVDPDNEAYEMPSEEGYQDY EPEA.

Full length α -synuclein was purified according to published protocols³⁴. The α -synuclein construct was transformed into *Escherichia coli* expression cell line BL21 (DE3) gold (Agilent Technologies) for wild-type α -synuclein protein expression. A single colony was incubated into 100 ml LB Miller broth (Fisher Scientific) supplemented with $100 \mu\text{g ml}^{-1}$ ampicillin (Fisher Scientific) and grown overnight at 37°C . One litre of LB (Miller) supplemented with $100 \mu\text{g ml}^{-1}$ ampicillin in 2-l shaker flasks was incubated with 10 ml of overnight culture and grown at 37°C until the culture reached $\text{OD}_{600} \approx 0.6\text{--}0.8$ as measured by a BioPhotometer UV/VIS Photometer (Eppendorf). IPTG (Isopropyl β -D-1-thiogalactopyranoside) was added to a final concentration of 0.5 mM, and grown for 4–6 h at 30°C . Cells were harvested by centrifugation at 5,500g for 10 min at 4°C . The cell pellet was frozen and stored at -80°C .

The cell pellet was thawed on ice and resuspended in lysis buffer (100 mM Tris-HCl pH 8.0, 500 mM NaCl, 1 mM EDTA pH 8.0) and lysed by sonication. Crude cell lysate was clarified by centrifugation at 15,000g for 30 min at 4°C . The clarified cell lysate was boiled and cell debris was removed by centrifugation. Protein in the supernatant was precipitated in acid at pH 3.5 through addition of HCl by titration to protein solution on ice while stirring then centrifuged for an additional 15,000g for 30 min at 4°C . Supernatant was dialysed against buffer A (20 mM Tris-HCl, pH 8.0). After dialysis the solution was filtered through a $0.45 \mu\text{m}$ syringe (Corning) before loading onto a 20 ml HiPrep Q HP 16/10 column (GE Healthcare). The Q-HP column was washed with five column volumes of buffer A and protein eluted using a linear gradient to 100% in five column volumes of buffer B (20 mM Tris-HCl, 1 M NaCl, pH 8.0). Protein eluted at around 50–70% buffer B; peak fractions were pooled. Pooled samples were concentrated approximately tenfold using Amicon Ultra-15 centrifugal filters. Approximately 5 ml of the concentrated sample was loaded onto a HiPrep 26/60 Sephacryl S-75 HR column equilibrated with filtration buffer (25 mM sodium phosphate, 100 mM NaCl, pH 7.5). Peak fractions were pooled from the gel filtration column and dialysed against 5 mM Tris-HCl, pH 7.5, concentrated to 3 mg ml^{-1} . These were filtered through a $0.2 \mu\text{m}$ pore size filter (Corning) and stored at 4°C .

Recombinantly expressed full-length α -synuclein with an N-terminal acetylation was prepared and purified in the following way based on a protocol detailed in ref. 16 The α -synuclein plasmid was co-expressed with a heterodimeric protein acetylation complex from *Schizosaccharomyces pombe* to acetylate the N terminus (pACYC-DUET, chloramphenicol, T7 promoter)⁵⁵. The two vectors were co-transformed into *E. coli* BL21 (DE3) using media containing both ampicillin and chloramphenicol. Cell cultures were grown in TB media containing ampicillin and chloramphenicol and induced to express α -synuclein with 0.5 mM IPTG overnight at 25°C . Cells were harvested by centrifugation, the cell pellet then resuspended in lysis buffer (100 mM Tris-HCl pH 8.0, 500 mM NaCl, 1 mM EDTA pH 8.0, and 1 mM phenylmethylsulfonyl fluoride) and cells lysed using an Emulsiflex homogenizer (Avestin). The lysate was boiled and debris removed by centrifugation. A protein fraction was also removed by precipitation at low pH on ice followed by centrifugation. The remaining supernatant was pH adjusted by titration and dialysed against buffer A (20 mM Tris-HCl, pH 8.0,

1 mM DTT, 1 mM EDTA, pH 8.0). The resulting protein solution was loaded onto a 5 ml Q-Sepharose FF column (GE Healthcare) equilibrated with buffer A and eluted against a linear gradient of buffer B (1 M NaCl, 20 mM Tris-HCl, pH 8.0, 1 mM DTT, 1 mM EDTA, pH 8.0). Fractions containing α -synuclein were identified using SDS-PAGE, collected, concentrated and further purified by size exclusion (Sephacryl S-100 16/60, GE Healthcare) in 20 mM Tris, pH 8.0, 100 mM NaCl, 1 mM DTT, 1 mM EDTA. Purity of fractions was assessed by SDS-PAGE.

Acetylated protein was characterized by LC-MS^{30,56}. Expected average mass: 14460.1 Da for α -synuclein and 14502.1 Da for acetylated α -synuclein. Observed average mass: 14464.0 Da for α -synuclein and 14506.0 Da for acetylated α -synuclein (Extended Data Fig. 6). The shift of 4 Da between observed and expected average masses is due to instrumental error.

Fibril formation and detection. Purified α -synuclein in 50 mM Tris, 150 mM KCl, pH 7.5 was shaken at a concentration of 500 μ M at 37 °C in a Torey Pine shaker. To form the fibrillar samples of SubNACore and NACore, lyophilized peptides were dissolved to a final concentration of 500 μ M in 5 mM lithium hydroxide, 20 mM sodium phosphate pH 7.5 and 0.1 M NaCl. All samples were shaken at 37 °C in a Torey Pine shaker for 72 h. Freshly dissolved samples were prepared by dissolving lyophilized peptides immediately before addition to cells for assays.

Turbidity measurements were used to compare NACore and SubNACore aggregation. Peptide samples were freshly dissolved to 1.6 mM in a sample buffer with 5 mM LiOH and 1% DMSO and then filtered through a PVDF filter (Millipore, 0.1 μ m). Measurements were performed using a black NUNC 96 well plate with 200 μ l of sample per well (3–4 replicates per sample). The plate was agitated at 37 °C, with a 3 mm rotation diameter, at 300 r.p.m. in a Varioskan microplate reader (Thermo). Absorbance readings were recorded every 3–15 min at 340 nm.

Negative-stain transmission electron microscopy. Cytotoxicity samples were evaluated for presence of fibrils by electron microscopy. In brief, 5- μ l samples were spotted directly on freshly glow-discharged carbon-coated electron microscopy grids (Ted Pella). After 4 min incubation, grids were rinsed twice with 5 μ l distilled water and stained with 2% uranyl acetate for 1 min. Specimens were examined on an FEI T12 electron microscope.

Fibril diffraction. Fibrils formed from purified α -synuclein with and without N-terminal acetylation were concentrated by centrifugation, washed, and oriented while drying between two glass capillaries. Likewise, NACore nanocrystals were also concentrated, washed with nanopure water, and allowed to orient while drying between two glass capillaries. The glass capillaries holding the aligned fibrils or nanocrystals were mounted on a brass pin for diffraction at room temperature using 1.54 Å X-rays produced by a Rigaku FRE+ rotating anode generator equipped with an HTC imaging plate. All patterns were collected at a distance of 180 mm and analysed using the Adxv software package⁵⁷. A simulated pattern from the full length α -synuclein model presented in Fig. 3 was obtained by calculating structure factors from the model using the sfall module from CCP4, assigning the model a unit cell of $200 \times 4.74 \times 200$ Å. Cylindrical averaging of these structure factors about the fibre axis (y axis) direction produced a set of simulated fibril diffraction intensities.

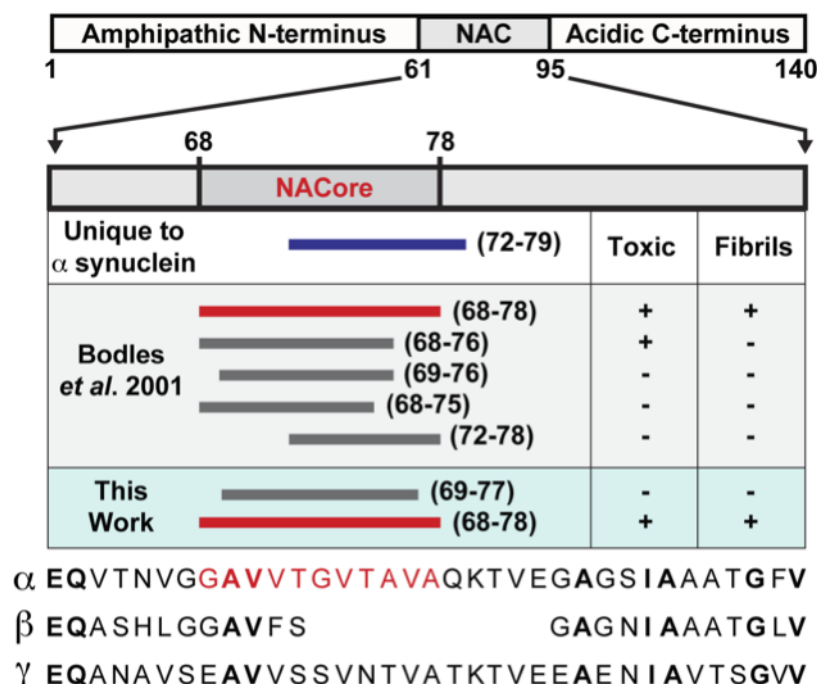
Cytotoxicity assays. Adherent PC12 cells (ATCC CRL-1721) were cultured in ATCC-formulated RPMI-1640 medium (ATCC 30-2001) supplemented with 10% horse serum and 5% fetal bovine serum and plated at 10,000 cells per well to a final volume of 90 μ l. All MTT assays were performed with Cell Titer 96 aqueous non-radioactive cell proliferation kit (MTT, Promega cat. no. 4100). Cells were cultured in 96-well plates for 20 h at 37 °C in 5% CO₂ before addition of samples (Costar cat. no. 3596). 10 μ l of sample was added to each well containing 90 μ l of medium and incubated for 24 h at 37 °C in 5% CO₂. Then, 15 μ l dye solution (Promega cat. no. 4102) was added into each well, followed by incubation for 4 h at 37 °C in 5% CO₂. This was followed by the addition of 100 μ l solubilization Solution/Stop Mix (Promega cat. no. 4101) to each well. After 12 h incubation at room temperature, the absorbance was measured at 570 nm. Background absorbance was recorded at 700 nm. The data was normalized with cells treated with 1% (w/v) SDS to 0% reduction, and cells treated with sample buffer to 100% reduction.

Lactose dehydrogenase assays were done using CytoTox-ONE Homogeneous Membrane Integrity, (Promega, cat. no. G7890) as per manufacturer's instructions.

In brief, cells were plated in 96-well, black-wall, clear-bottom (Fisher cat. no. 07-200-588) tissue culture plates at 10,000 cells per well to a final volume of 90 μ l. Cells were incubated for an additional 20 h at 37 °C in 5% CO₂ before addition of samples. Next, 10 μ l of sample was added to each well, following which the cells were incubated for another 24 h. 100 μ l of reagent was added to each well and incubated for 15 min at room temperature. The addition of 50 μ l of stop solution stopped the reaction. Fluorescence was measured in a Spectramax M5 (Molecular Devices) using excitation and emission wavelengths of 560 nm and 590 nm, respectively. Data was normalized using cells treated with buffer as 0% release and 0.1% Triton X-100 as 100% release.

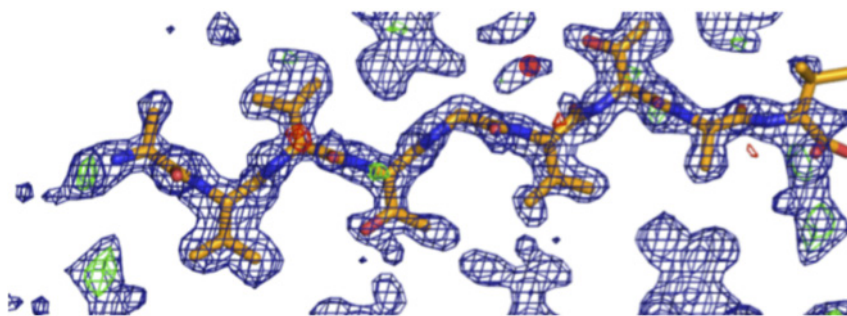
Construction of α -synuclein A53T fibril model. A model for full-length α -synuclein A53T mutant fibrils that are involved in the early onset of PD was constructed using a section of the NACore crystal packing as a scaffold. Figure 3 illustrates the four copies of the NACore segment used for the scaffold. The crystal structure of the two inner strands was adapted with minimal changes as the analogous segments 68–78. The structure of PreNAC was matched onto the weak interface of the NACore structure. Only 4 of the 11 side chains in the segment 46–56 differ from those in the NACore segment 68–78 and residues V51–V55 can be closely matched to V71–V75. Hence the model for both the homotypic interface and heterotypic interface in the full-length fibre model closely resemble those observed in the NACore structure. The regions outside these segments were adapted from the structure of the native α -synuclein fold, PDB ID: 2KKW⁵⁸. These segments were spliced in manually using COOT. The models were energy minimized and temperature annealed using the program CNS⁵⁹ with hydrogen-bonding potential⁶⁰. The simulated fibre diffraction pattern calculated from this model shows prominent reflections that agree with those observed in fibre diffraction patterns of NACore, α -synuclein, and N-acetyl α -synuclein (Extended Data Table 2).

43. Otwinowski, Z. & Minor, W. in *Methods in Enzymology* (eds Carter, C. W. Jr & Sweet, R. M.) Vol. 276 307–326 (Academic Press, 1997).
44. Weierstall, U., Spence, J. C. H. & Doak, R. B. Injector for scattering measurements on fully solvated biospecies. *Rev. Sci. Instrum.* **83**, 035108 (2012).
45. Hattne, J. *et al.* Accurate macromolecular structures using minimal measurements from X-ray free-electron lasers. *Nature Methods* **11**, 545–548 (2014).
46. Sauter, N. K., Hattne, J., Grosse-Kunstleve, R. W. & Echols, N. New Python-based methods for data processing. *Acta Crystallogr. D* **69**, 1274–1282 (2013).
47. Kabsch, W. XDS. *Acta Crystallogr. D* **66**, 125–132 (2010).
48. McCoy, A. J. *et al.* Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674 (2007).
49. Murshudov, G. N., Vagin, A. A. & Dodson, E. J. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D* **53**, 240–255 (1997).
50. Afonine, P. V. *et al.* Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr. D* **68**, 352–367 (2012).
51. Blanc, E. *et al.* Refinement of severely incomplete structures with maximum likelihood in BUSTER-TNT. *Acta Crystallogr. D* **60**, 2210–2221 (2004).
52. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D* **66**, 486–501 (2010).
53. Delano, W. *The PyMOL Molecular Graphics System* (Schrödinger LLC) <http://www.pymol.org>.
54. Jakes, R., Spillantini, M. G. & Goedert, M. Identification of two distinct synucleins from human brain. *FEBS Lett.* **345**, 27–32 (1994).
55. Johnson, M., Coulton, A. T., Gees, A. & Mulvihill, D. P. Targeted amino-terminal acetylation of recombinant proteins in *E. coli*. *PLoS ONE* **5**, e15801 (2010).
56. Whitelegge, J. P., Zhang, H., Aguilera, R., Taylor, R. M. & Cramer, W. A. Full subunit coverage liquid chromatography electrospray ionization mass spectrometry (LCMS+) of an oligomeric membrane protein: cytochrome *b₆f* complex from spinach and the cyanobacterium *Mastigocladus laminosus*. *Mol. Cell. Proteomics MCP* **1**, 816–827 (2002).
57. Arvai, A. *Adxv - A Program to Display X-ray Diffraction Images* (2015).
58. Rao, J. N., Jao, C. C., Hegde, B. G., Langen, R. & Ulmer, T. S. A combinatorial NMR and EPR approach for evaluating the structural ensemble of partially folded proteins. *J. Am. Chem. Soc.* **132**, 8657–8668 (2010).
59. Brunger, A. T. Version 1.2 of the crystallography and NMR system. *Nature Protocols* **2**, 2728–2733 (2007).
60. Fabiola, F., Bertram, R., Korostelev, A. & Chapman, M. S. An improved hydrogen bond potential: impact on medium resolution protein structures. *Protein Sci.* **11**, 1415–1423 (2002).
61. Read, R. J. Improved Fourier coefficients for maps using phases from partial structures with errors. *Acta Crystallogr. A* **42**, 140–149 (1986).



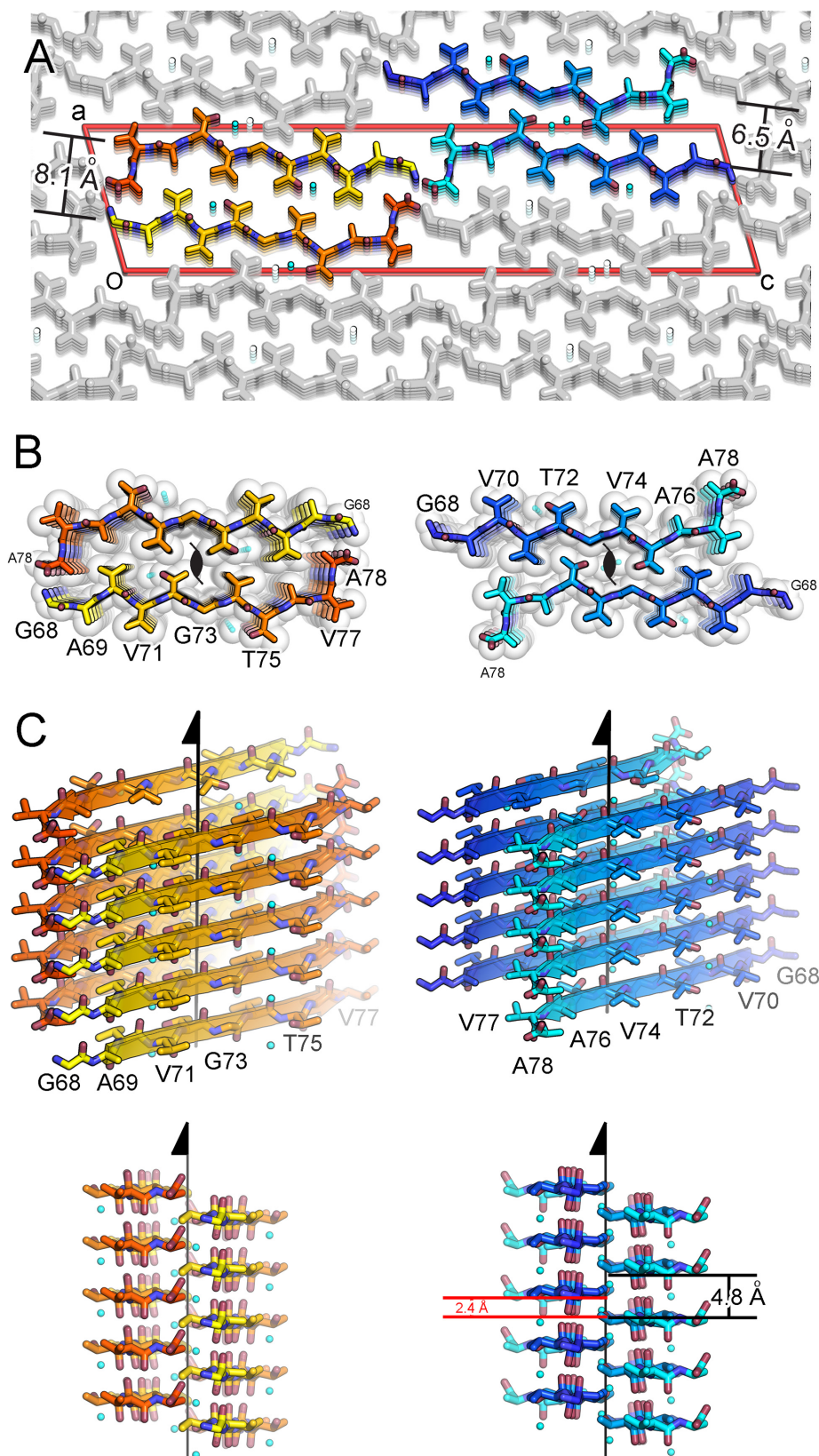
Extended Data Figure 1 | A schematic representation of α -synuclein, highlighting the NAC region (residues 61–95) and within it the NACore sequence (residues 68–78). A series of bars span regions of α -synuclein that are of interest to this work. Among the three synuclein paralogues (α , β and γ), the region whose sequence is unique to α -synuclein is shown as a blue bar (residues 72–83) that overlaps with a large portion of NACore. Segments investigated in ref. 23 are also shown. These span a variety of regions within

NACore. Two of the segments we investigate here, SubNACore and NACore, are shown in this context. Only one of the segments studied ref. 23 is an exact match to our NACore sequence, and only this segment is both toxic and fibrillar. The sequences of α -synuclein, β -synuclein, and γ -synuclein are shown as a reference with conserved residues in bold and the NACore sequence in red.



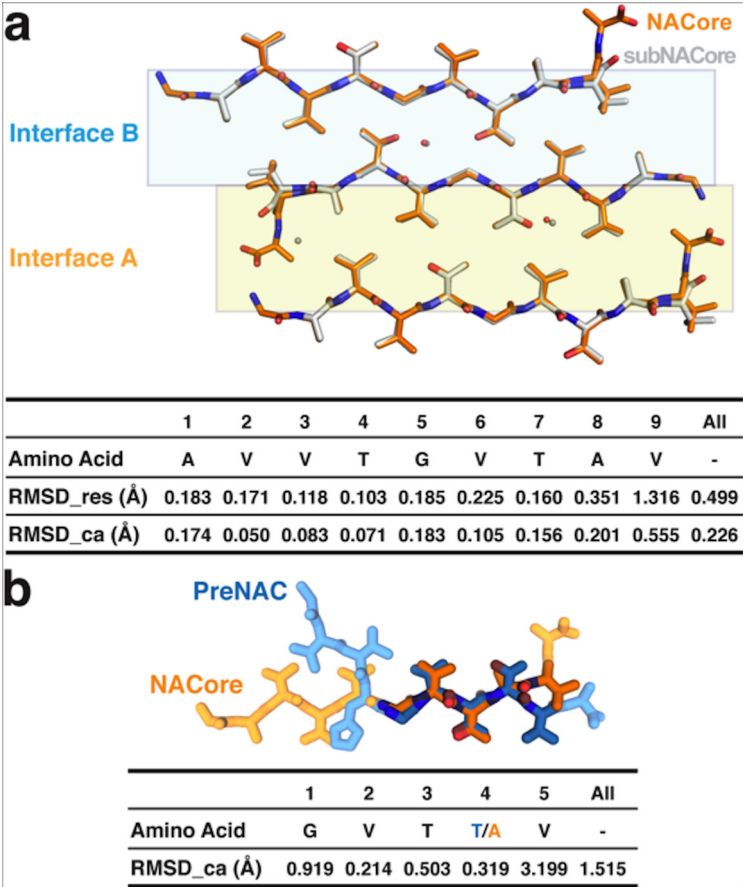
Extended Data Figure 2 | NACore difference density maps calculated after successful molecular replacement using the SubNACore search model clearly revealed the positions of the missing residues (positive $F_o - F_c$ density at N and C termini corresponding to G68 and A78) and one water

molecule near a threonine side chain (red circle); a second water was located during the refinement process. The blue mesh represents $2F_o - F_c$ density contoured at 1.2σ . The green and red mesh represent $F_o - F_c$ densities contoured at 3.0 and -3.0σ , respectively. All maps were σ_A -weighted⁶¹.



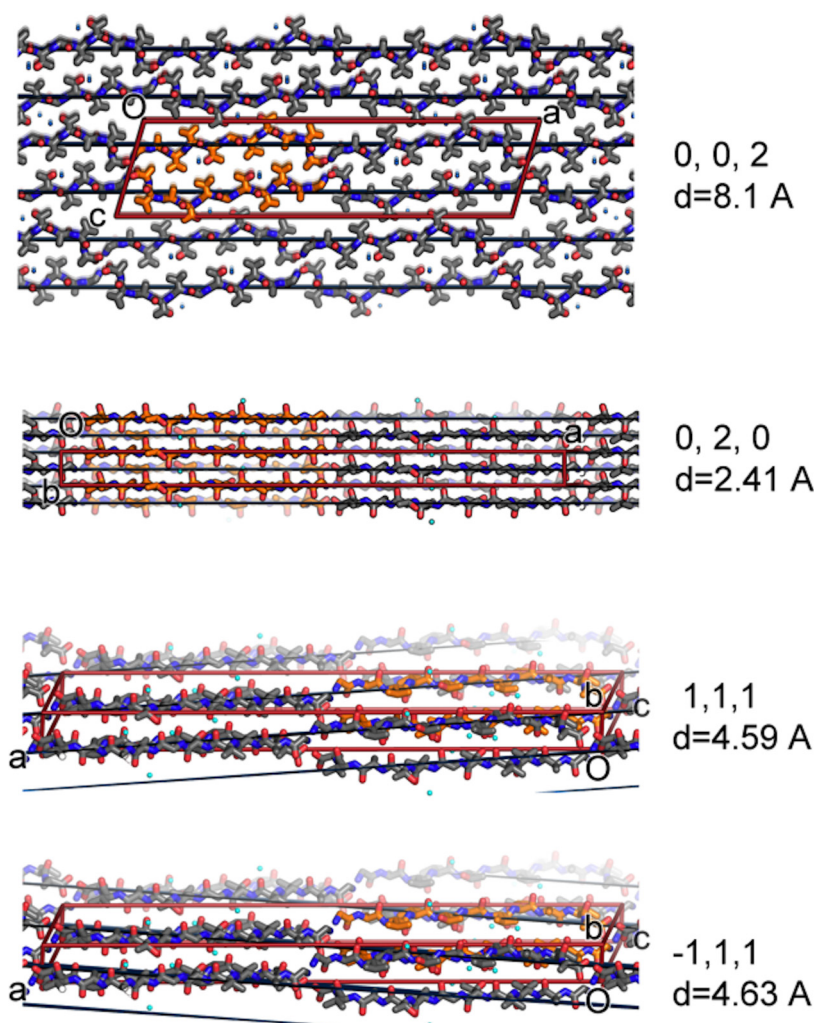
Extended Data Figure 3 | The crystal structure of NACore reveals pairs of sheets as in the spines of amyloid fibrils. **a**, NACore's two types of sheet-sheet interfaces: a larger interface (orange, 268 \AA^2 of buried accessible surface area per chain) we call interface A, and a weaker interface (blue, 167 \AA^2) we call interface B. The crystal is viewed along the hydrogen-bonding direction (crystal 'b' dimension). The red lines outline the unit cell. **b**, The van der Waals packing between sheets. The sheets are related by a 2_1 screw axis denoted in

black. The only gaps left by the interface are filled with water molecules which hydrogen-bond to the threonine residues (partially showing aqua spheres). The shape complementarity of both interfaces is 0.7. The viewing direction is the same as in **a**. **c**, Orthogonal view of the fibrillar assembly. The protofibril axis, coinciding with the 2_1 screw axis designated by the arrow, runs vertically between the pairs of sheets.



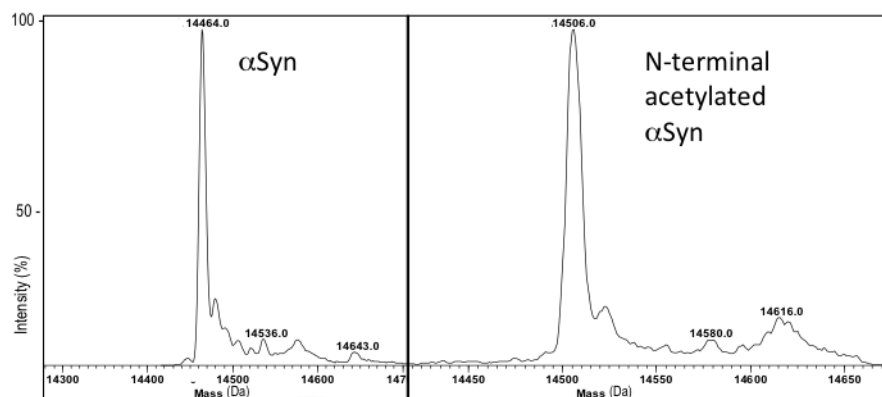
Extended Data Figure 4 | Comparison of the crystal packing for NACore and SubNACore. **a**, The face-to-face interactions are virtually the same for the pairs of NACore segments (orange chains) in its crystal structure and the SubNACore segments (white chains) in its structure (interfaces A and B shown in gold and blue, respectively). The table below shows the pairwise r.m.s.d. values comparing the nine residues shared in common between the structures.

RMSD_res is an all-atom comparison between residue pairs, while RMSD_ca compares only C α pairs. **b**, PreNAC (blue) is compared with NACore (orange). Five residues from each strand are shown in darker colour and the r.m.s.d. values between their C α pairs are compared in the table below. The PreNAC–NACore interaction mimics the weaker interface B in the NACore structure.



Extended Data Figure 5 | Intense reflections common among the NACore and the two polymorphs of full length α -synuclein suggest common structural features. Common structural features are illustrated here on the crystal packing diagrams of NACore. The $(0,0,2)$ planes approximate the separation between sheets in interface A (orange). The $(0,2,0)$, $(-1,1,1)$, and $(1,1,1)$ reflections are intense because the corresponding Bragg planes

recapitulate the staggering of strands from opposing sheets. The red lines correspond to the unit cell boundaries and all planes are shown in black. The location of the unit cell origin is indicated by 'O'. The unit cell dimensions a , b , and c are labelled. Bragg spacings (spacings between planes), indicated by 'd', are indicated in angstroms.



Extended Data Figure 6 | Mass spectrometry analysis of recombinantly expressed, full-length α -synuclein, with and without N-terminal acetylation. The mass profile of wild-type full length α -synuclein (left) is compared to that of an N-terminally acetylated form of the protein (right). The mass shift for

the N-terminally acetylated form is appropriately shifted with respect to the native form of the protein (14464.0 Da for α -synuclein and 14506.0 Da for acetylated α -synuclein), within a margin of error of 4 Da.

Extended Data Table 1 | Statistics of data collection and atomic refinement for NACore, its fragment SubNACore, and PreNAC

Segment	SubNACore 69AVVTGVTAV ₇₇	NACore 68GAVVTGVTAVA ₅₆	PreNAC 47GVVHGVTTVA ₅₆
Data collection			
Radiation source	Synchrotron	Electron	Electron
Space group	C2	C2	P21
Cell dimensions			
<i>a, b, c</i> (Å)	61.9, 4.80, 17.3	70.8, 4.82, 16.79	17.9, 4.7, 33.0
α, β, γ (°)	90, 104.1, 90	90, 105.7, 90	90, 94.3, 90
Resolution (Å)	1.85 (1.95-1.85)	1.43 (1.60-1.43)	1.41 (1.56-1.41)
Wavelength (Å)	0.9791	0.0251	0.0251
<i>R</i> _{merge}	0.117 (0.282)	0.173 (0.560)	0.236 (0.535)
<i>R</i> _{r.i.m.}	0.135 (0.322)	0.199 (0.647)	0.264 (0.609)
<i>R</i> _{p.i.m.}	0.065 (0.154)	0.093 (0.311)	0.185 (0.305)
<i>I</i> / σ <i>I</i>	5.2 (2.7)	5.5 (2.5)	4.6 (1.8)
CC _{1/2} (%)	99.5 (97.8)	99.4 (92.3)	96.7(74.0)
Completeness (%)	97.9 (98.3)	89.9 (82.6)	86.9 (69.6)
Multiplicity	4.1 (4.0)	4.4 (4.3)	3.7 (3.5)
Refinement			
Resolution (Å)	1.85 (2.07-1.85)	1.43 (1.60-1.43)	1.41 (1.41-1.57)
No. reflections	470 (125)	1073 (245)	1006 (239)
<i>R</i> _{work}	0.176 (0.248)	0.248 (0.253)	0.235 (0.336)
<i>R</i> _{free}	0.221 (0.286)	0.275 (0.331)	0.282 (0.329)
CC _{work}	0.964 (0.896)	0.947(0.618)	0.937(0.335)
CC _{free}	0.889 (0.993)	0.986(0.269)	0.967(0.361)
No. atoms			
Protein	57	66	66
Water	3	2	4
B-factors (Å ²)			
Protein	17.1	9.0	16.1
Water	27.6	2.7	24.6
Wilson B (Å ²)	11.8	10.3	13.8
R.m.s deviations			
Bond lengths (Å)	0.005	0.010	0.020
Bond angles (°)	1.1	1.6	2.0
PDB ID code	4RIK	4RIL	4ZNN
EMDB ID code		EMD-3028	EMD-3001

*Highest resolution shell is shown in parenthesis. Data quality is indicated by the redundancy independent merging *R* factor (r.i.m) and the precision indicating merging *R* factor (p.i.m.).

Extended Data Table 2 | Comparison of reflections observed in powder diffraction of fibrils of full-length α -synuclein, N-acetyl α -synuclein, and a synthetic pattern calculated from our α -synuclein model, to aligned nanocrystals of NACore.

Segment	Reflections (Å)
NACore GAVVTGVTAVA	2.21, 2.26, 2.39 , 2.52, 2.61, 2.68, 2.78, 3.02, 3.12, 3.34, 3.56, 3.86, 4.34, 4.57 , 5.16, 5.98, 7.56, 8.19 , 10.46, 11.63, 13.29, 16.61
α -syn	2.39 , 4.64 , 6.82, 8.29 , 10.06
N-acetyl α -syn	2.38 , 4.62 , 8.18 , 9.80, 11.90
Simulated α -syn	2.23, 2.25, 2.35 , 3.29, 3.63, 3.70, 3.95, 4.08, 4.56, 4.68 , 8.36 , 8.69, 21.76, 24.47, 27.61, 31.67

Bold reflections are strong and common to all three samples. Colours of the labelled reflections match those in Fig. 2.

Structure of mammalian eIF3 in the context of the 43S preinitiation complex

Amedee des Georges¹, Vidya Dhote², Lauriane Kuhn³, Christopher U. T. Hellen², Tatyana V. Pestova², Joachim Frank^{1,4} & Yaser Hashem⁵

During eukaryotic translation initiation, 43S complexes, comprising a 40S ribosomal subunit, initiator transfer RNA and initiation factors (eIF) 2, 3, 1 and 1A, attach to the 5'-terminal region of messenger RNA and scan along it to the initiation codon. Scanning on structured mRNAs also requires the DEXH-box protein DHX29. Mammalian eIF3 contains 13 subunits and participates in nearly all steps of translation initiation. Eight subunits having PCI (proteasome, COP9 signalosome, eIF3) or MPN (Mpr1, Pad1, amino-terminal) domains constitute the structural core of eIF3, to which five peripheral subunits are flexibly linked. Here we present a cryo-electron microscopy structure of eIF3 in the context of the DHX29-bound 43S complex, showing the PCI/MPN core at ~6 Å resolution. It reveals the organization of the individual subunits and their interactions with components of the 43S complex. We were able to build near-complete polyaniline-level models of the eIF3 PCI/MPN core and of two peripheral subunits. The implications for understanding mRNA ribosomal attachment and scanning are discussed.

Translation initiation in eukaryotes begins with binding of eIF3, eIF1, eIF1A and the eIF2-GTP-Met-tRNA^{Met} ternary complex to the 40S subunit, forming a 43S preinitiation complex¹. The 43S complex attaches to the cap-proximal region of mRNA after unwinding of its secondary structure by eIF4A, eIF4B and eIF4F, and scans downstream to the initiation codon, where it forms a 48S initiation complex by codon-anticodon base pairing. Scanning on structured mammalian mRNAs also requires DHX29, which binds directly to the 40S subunit. Finally, eIF5 and eIF5B promote joining of the 60S subunit to the 48S complex, yielding an elongation-competent 80S ribosome.

eIF3 is the largest, most complex initiation factor, which interacts with several eIFs, including eIF1 and the eIF4G subunit of eIF4F (refs 1–3). eIF3 is involved in almost all steps of initiation, including ribosomal recruitment of the ternary complex, attachment of 43S complexes to mRNA via interaction with eIF4G, and scanning. The ~800 kilodalton (kDa) mammalian eIF3 comprises 13 subunits (a–m) (Extended Data Fig. 1a). Six subunits (a, c, e, k, l and m) contain PCI domains, which consist of N-terminal helical repeats followed by a winged helix domain (WHD) that mediates PCI polymerization⁴, and two subunits (f and h) contain MPN domains, which consist of a β -barrel surrounded by α -helices and β -strands that function to promote assembly of multiprotein complexes^{5,6}. The PCI/MPN subunits form the octameric structural core of eIF3. Cryo-electron microscopy (cryo-EM) studies^{7–10} revealed the organization of the five-lobed PCI/MPN core of mammalian eIF3, and confirmed the similarity of its topology with those of the proteasome lid and the COP9 signalosome^{6,11–13}. However, the resolution of eIF3 in these studies (12–20 Å) was insufficient to reveal molecular details of the PCI/MPN core organization.

Four of the remaining subunits (b, d, g and i) are stably linked to the PCI/MPN core, probably in a flexible manner^{7,8,10}. Domains in these subunits include RNA recognition motif (RRM) domains (eIF3b and eIF3g)¹⁴ and WD40 β -propeller domains (eIF3b and eIF3i)^{15–17}. eIF3b, eIF3i and eIF3g form a separate module, which attaches to

the PCI/MPN core through its interaction with the eIF3a carboxy-terminal domain (CTD)^{15–21}. The last subunit, eIF3j, is substoichiometric and loosely attached to the rest of eIF3 (ref. 22). Whereas most eukaryotes encode a complete set of eIF3 subunits, *Saccharomyces cerevisiae* and related yeasts retain only six: two PCI (a and c) and four non-core (b, i, g and j) subunits, with eIF3j being non-essential²³.

We recently determined the structure of mammalian eIF3 and its position on the 40S subunit by cryo-EM reconstruction of the DHX29-bound 43S complex at 11.6 Å resolution⁷. The PCI/MPN core resides on the back of the 40S subunit, making two contact points via its left arm and head with ribosomal proteins (rp) eS1/eS26 and uS15/eS27, respectively. Two additional densities, on the solvent side underneath helix 16 (h16) and on the head behind RACK1, were attributed to the peripheral domains of eIF3 belonging to non-core subunits. However, the resolution of this complex was insufficient for modelling of eIF3. Thus, although recent crystallographic studies revealed some important structural aspects of yeast eIF3 (ref. 17), molecular details of mammalian eIF3 organization remained obscure.

Here we present a high-resolution cryo-EM reconstruction of mammalian eIF3 (lacking eIF3j) in the context of the DHX29-bound 43S complex. The reconstructed density map allowed us to derive a near-complete polyaniline-level model of the eIF3 PCI/MPN octamer core and two peripheral subunits, using homology and *ab initio* modelling.

Sample preparation and electron microscopy

To obtain the structure of 40S-associated eIF3, DHX29-bound 43S complexes were prepared as described⁷ (Supplementary Information). eIF3 was purified from rabbit reticulocyte lysate and contained C-terminally truncated eIF3a (2–1061; see Supplementary Information ‘Sample composition and image processing’, Extended Data Fig. 2a and Extended Data Table 1a).

The processed imaged particles yielded an ~6 Å cryo-EM reconstruction on average (Extended Data Figs 1b and 2d, f). Local refinement of the

¹HHMI, Department of Biochemistry and Molecular Biophysics, Columbia University, New York, New York 10032, USA. ²Department of Cell Biology, SUNY Downstate Medical Center, Brooklyn, New York 11203, USA. ³CNRS, Proteomic Platform Strasbourg - Esplanade, Strasbourg 67084, France. ⁴Department of Biological Sciences, Columbia University, New York, New York 10032, USA. ⁵CNRS, Architecture et Réactivité de l'ARN, Université de Strasbourg, Strasbourg 67084, France.

orientations of the images was applied to improve the resolution of eIF3 (Extended Data Fig. 2c, e) leading to a resolution of ~ 6 Å for most regions of the eIF3 core. Compared to our previous study⁷, secondary structure elements are now clearly resolved in most regions of the core (Extended Data Fig. 1c–f).

Structure and atomic model of eIF3 PCI/MPN core

The PCI/MPN octamer core adopts the classic five-lobed shape. Density segmentation (Fig. 1a–c) was performed on the basis of topological similarity with the proteasome lid and the COP9 signalosome^{11,13}, and rigid-body fitting of crystal structures of yeast eIF3a and eIF3c (ref. 17). This analysis revealed the general organization of the eIF3 core (Fig. 1d). The PCI subunits (a, c, e, l, k and m) are arranged sequentially, forming an arc. The MPN subunits (f and h) bind to each other and attach to the rest of the octamer mainly through the association of eIF3f with eIF3m. In addition to the PCI arc, at least one α -helix of each subunit (except for eIF3a and eIF3m) is involved in the formation of a seven-helix bundle.

Each subunit of the eIF3 core was modelled in its density segment using *ab initio* and homology modelling. When there was no homology to rely on, the secondary structure elements (principally α -helices) of the concerned parts were predicted and built into the map. The resulting model fits the eIF3 core density (Figs 1e–g and 2) with a cross-correlation coefficient of 0.94. Despite the good resolution of the eIF3 core in the cryo-EM map, one region of density could not be assigned (Extended Data Fig. 1g).

The model of the eIF3 core reveals that the PCI arc, the major interaction hub, consists of a large arched β -sheet assembled from the WHD β -sheets of the individual PCI domains of subunits a, c, e, l, k and m (Fig. 2c). This first interaction hub wraps around the second, which comprises a seven-helix bundle formed by packing of the C-terminal helices of subunits c, e, f, h, k and l (Fig. 2a, c, g). To assign helices to individual subunits, we used the conserved helical bundle structure in the proteasome lid and COP9 signalosome^{13,24}, as the

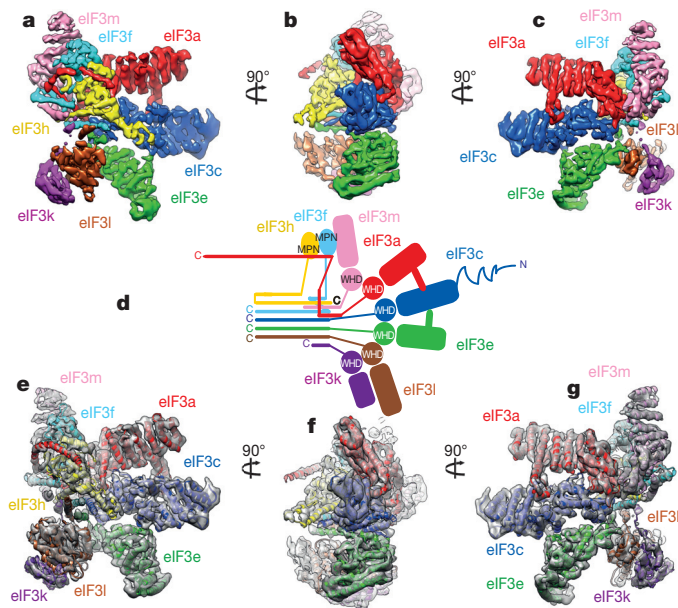


Figure 1 | Structure of eIF3 core. a–c, Segmented eIF3 core, coloured variably by subunit, seen from different orientations. d, Two-dimensional representation of the three-dimensional structure of the eIF3 octamer core. The helical bundle is represented by coloured bars. Zig-zagged line on eIF3c indicates possibly unstructured N-terminal tail. The eIF3a C-terminal region (not present in the structure reported here) is not shown in this schematic representation. e–g, Fitting of the eIF3 core model in its cryo-EM segmented density.

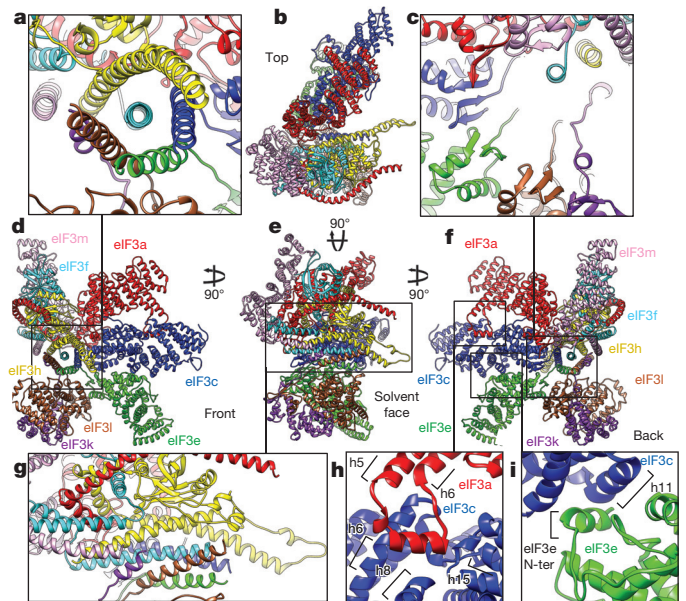


Figure 2 | Model of the eIF3 core. a, g, Close-up view of the seven-helix bundle formed by subunits h, c, e, l and k, seen in two orientations. b–f, Polyaniline-level model of the eIF3 octamer core seen from different orientations. c, Close-up view of β -sheet arc of the PCI domains of subunits m, a, c, e, k and l. The seven-helix bundle was cut out by fading it to highlight the arched β -sheet. h, i, Close-up views of additional quaternary contacts between eIF3a and eIF3c (h) and eIF3c and eIF3e (i).

organization of the seven-helix bundle is conserved in eIF3 (Extended Data Fig. 3a–c).

A comparison of our model of the PCI/MPN core of eIF3 with those of the proteasome lid and the COP9 signalosome (Extended Data Fig. 3d–i) showed the existence of additional quaternary interactions between eIF3a and eIF3c, and between eIF3c and eIF3e (Fig. 2h, i and Extended Data Fig. 3j, k). eIF3a interacts through an insert between helices 5 and 6 with eIF3c in a cavity formed by helix 6, the coil between helices 7 and 8 and the coil between helices 14 and 15. The eIF3c–eIF3e interaction involves helix 11 of eIF3c and the N-terminal tail of eIF3e. These interactions rigidify the assembly of eIF3a, eIF3c and eIF3e, and the rotation of their helical-repeat domains with respect to their WHDs (Extended Data Fig. 3l). Notably, the insert in mammalian eIF3a, which participates in the quaternary interaction, is lacking in yeast eIF3a. Mammalian eIF3a also possesses a more complex C-terminal helix following the PCI domain, which is involved in additional interactions with other core subunits (Figs 1d and 2).

To validate our model, amino acid conservation at the interfaces of core subunits was assessed among five multicellular (*Homo sapiens*, *Caenorhabditis elegans*, *Xenopus tropicalis*, *Arabidopsis thaliana* and *Drosophila melanogaster*) and one unicellular (*Neurospora crassa*) organisms containing eIF3 with similar subunit composition (Supplementary Information ‘Validation of eIF3 core model’ and Extended Data Figs 4–6). Importantly, the a–c and c–e interfaces involving the WHDs are conserved in multicellular organisms analysed, but not in *N. crassa*. However, the additional quaternary interactions between these subunits are conserved even in *N. crassa*, suggesting their importance in all eukaryotes that retain the octameric composition of the eIF3 core. The interfaces between the remaining subunits are conserved among all analysed organisms with few exceptions in *N. crassa*.

Our current model exhibits large discrepancies with a recent model¹⁷ based on low-resolution data concerning the structure and conformation of various core subunits (that is, differences in the conformations of eIF3f, eIF3h and eIF3e, as well as in the structures

of the C-terminal α -helices of eIF3f and eIF3h, the assignment of which was inverted in that model, thereby altering the structure of the seven-helix bundle) (Extended Data Fig. 7).

Ribosomal contacts of the PCI/MPN core

To identify residues in eIF3 core that directly contact components of the 43S complex, the atomic model of the human 40S subunit²⁵, the crystal structure of the archaeal eIF2–ternary complex²⁶, and the homology model of the C-terminal two-thirds of DHX29 (ref. 27) were fitted into the density (Methods). Direct interactions of eIF3a and eIF3c with the 40S subunit were observed. eIF3a contacts rpeS1 and, marginally, rpeS26 (Fig. 3a, pink and yellow arrowheads), but also the apical loop of ES7^S (Fig. 3, top grey arrowhead). eIF3c interacts with rpeS27 and, marginally, with rpuS15 (Fig. 3, green and cyan arrowheads), but also with the apical loop of ES7^S (Fig. 3, bottom grey arrowhead) and h22 (Fig. 3, red arrowhead). A detailed list of interactions is presented in Extended Data Table 1b. They are consistent with those observed for *S. cerevisiae* eIF3a and eIF3c (ref. 17), with the exception of the interaction between the C terminus of yeast eIF3a and rpuS2 (refs 28–30), which is absent in the mammalian complex. Furthermore, the yeast eIF3a–eIF3c core is rotated by $\sim 24^\circ$ away from the platform and towards the solvent side in contrast to mammalian eIF3 (ref. 28).

Peripheral subunits of eIF3

To improve the resolution in peripheral domains of eIF3, focused classification was applied to the regions where peripheral subunits are located (Extended Data Fig. 9a, b), yielding reconstructions with local resolutions of ~ 9 Å and ~ 10 Å for regions of eIF3 near h16 and on the head next to RACK1, respectively.

Segmentation of density corresponding to the peripheral domains of eIF3 near h16 revealed a nine-bladed β -propeller structure (Fig. 4a), unambiguously attributed to the WD40 domain of eIF3b (refs 16, 17). Density projecting from the WD40 domain towards the beak of the 40S subunit, in front of h16, corresponds to the eIF3b RRM domain. We have derived a homology model of residues 73–685 of rabbit eIF3b ($\sim 85\%$) (Fig. 4b). In the 43S complex, eIF3b binds through its C-terminal tail, after the WD40 domain (residues 626–631) to the 40S subunit at the tip of helix A (hA) of ES6^S (Fig. 4a, right, red arrowhead), and interacts with rpuS4 and DHX29 over a large surface (Fig. 4c, slate grey and green arrowheads) (Extended Data Table 1b). The position of eIF3b in mammalian DHX29-bound 43S complexes⁷ is similar to that in yeast 40S–eIF3–eIF1–eIF1A complexes¹⁷, despite the absence of DHX29 in the latter.

The low-resolution density extending from the eIF3b WD40 domain away from the 40S subunit (Fig. 4a, dashed oval) was attributed to eIF3i, on the basis of crystal structures of the latter with the C-terminal helix of yeast eIF3b (ref. 15) and of the yeast eIF3b–CTD–eIF3i–eIF3g–NTD complex¹⁷, and the size and shape compatibility with our segmented density. In the 43S complex, eIF3i is seen to

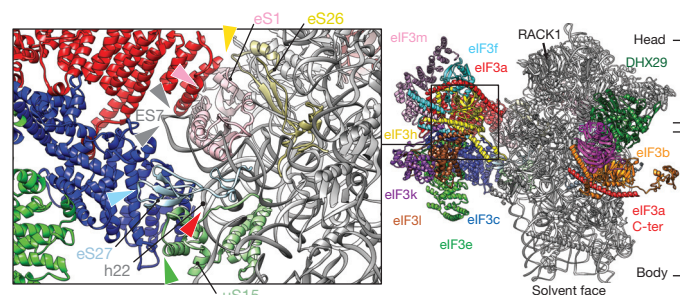


Figure 3 | Contacts of the eIF3 core with the 40S subunit in the 43S complex. Solvent-side view of eIF3 bound to the DHX29-bound 43S complex. In black rectangle, close-up view of 40S–eIF3 core contacts. Coloured arrowheads indicate the interaction of the eIF3 core with ribosomal proteins and RNA.

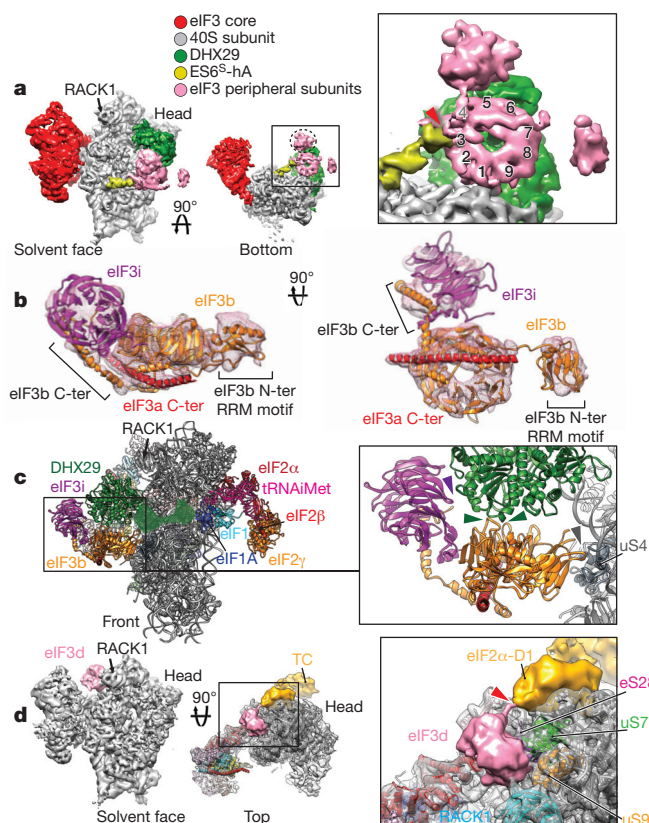


Figure 4 | Peripheral subunits of eIF3. **a**, Segmented cryo-EM reconstruction focused on eIF3 peripheral subunits localized near the mRNA channel entry, below DHX29, seen from the solvent face (left) and from below (middle). Close-up view of eIF3 peripheral subunits seen from below (right), with a red arrowhead indicating the connection with ES6^S-hA and numbers indicating the blades of the β -propeller structure. **b**, Atomic model of rabbit eIF3b (orange), yeast eIF3i (purple) and a long α -helix (red) corresponding to a fragment of the C-terminal helical region ('eIF3a C-ter'). **c**, Left, front view of eIF3b and eIF3i subunits bound to the 40S subunit and DHX29 in the context of the 43S complex. The remaining, unmodelled one-third of DHX29 is denoted as a transparent green surface, based on its cryo-EM density. Right, close-up view of the contact points between eIF3b and DHX29. The panel also shows the interaction of eIF3i with DHX29. **d**, Left, segmented cryo-EM reconstruction focused on an eIF3 peripheral subunit tentatively identified as eIF3d, localized on the 40S head behind ribosomal protein RACK1. Middle, the same reconstruction seen from above, rendered in transparent with the atomic model of the human 40S subunit fitted in. Right, close-up view of the putative eIF3d subunit. Red arrow indicates a density bridging the globular domain of eIF3d to a density corresponding to the eIF2 α -D1 domain, part of the eIF2-GTP-Met-trRNA^{Met} ternary-complex (TC) (density coloured in gold). Coloured arrowheads indicate the interaction of eIF3 peripheral subunits with various ribosomal proteins.

interact only with DHX29 (Fig. 4c, purple arrow). No density could be attributed to the eIF3g component of the eIF3g–eIF3b–eIF3i module because of the low local resolution in that region, probably due to the flexibility of this assembly.

The residues forming the long curved helix that interacts with eIF3b at its N and C termini (Fig. 4b, red helix) could not be determined with certainty because of its discontinuity with the eIF3 core. However, on the basis of its curvature and the angle it forms with the core, it seems to extend from the C-terminal helix of eIF3a. This attribution is consistent with observations that eIF3a, eIF3b, eIF3i and eIF3g form a stable complex^{15,17–21}, and with the existence of a spectrin domain at the eIF3a C-terminal end, which mediates interactions with eIF3b and eIF3i (ref. 19).

The second peripheral subunit-binding site, behind RACK1, displays a globular mass with a local resolution of ~ 10 Å and no resolved

secondary structure features (Fig. 4d and Extended Data Fig. 9a), which adopts a flat triangular shape at an increased density threshold. This mass was attributed to eIF3d, consistent with the latter's cross-linking with positions –8 to –17 on mRNA³¹ and because the only other unassigned subunit, eIF3g, constitutes part of the eIF3b–eIF3i–eIF3g module. According to our structure, eIF3d directly contacts RACK1, rpeS28, rpuS7 and rpuS9 (Fig. 4d, right, and Extended Data Fig. 9b). Notably, it also appears to interact with the D1 domain of eIF2 α , as seen at a lower density threshold (Fig. 4d, right, red arrow).

Discussion

Our reconstruction of mammalian eIF3 in the context of the DHX29-bound 43S complex enabled us to build a near-complete polyaniline-level model of the eIF3 octameric PCI/MPN core, and to determine the ribosomal positions of 11 of its 12 stoichiometrically associated subunits (Fig. 5). eIF3 forms three regions of density corresponding to the PCI/MPN core, the eIF3b–eIF3i–eIF3g module, and eIF3d. The PCI/MPN core binds at the solvent side of the 40S subunit opposite to the platform, with eIF3a and eIF3c establishing two contact points. The eIF3b–eIF3i–eIF3g module resides at the mRNA entrance (with eIF3b interacting directly with the 40S subunit), and is connected to the core by the eIF3a–CTD. eIF3d is located near the mRNA exit, behind RACK1.

The model of the PCI/MPN core reveals its assembly and intersubunit interactions in mammalian eIF3 (Figs 1 and 2), which are probably conserved in other organisms possessing eIF3 with a similar octameric core. This model is consistent with reports concerning the interactions of individual subunits and the existence of their stable sub-complexes^{2,9,10,15,17–21}, and with the finding that the functional core of mammalian eIF3 comprises one non-PCI/MPN subunit (b) and five PCI/MPN subunits (a, c, e, f and h)²¹. Thus, whereas eIF3a, eIF3c and eIF3e are the central constituents of the PCI arc, the PCI subunits k, l and m reside at its extremities (Fig. 1b, d). Consistently, eIF3k and eIF3l are easily displaced from eIF3 (ref. 18) and have been lost from the genome of some species³². eIF3m, which is not encoded in *Trypanosoma brucei* and *Leishmania*³³, can also be absent without major perturbation to the PCI arc. Importantly, the structure and arrangement of the ‘essential’ PCI subunits (a, c and e) are further rigidified by the quaternary a–c and c–e interactions, inexistent between their counterparts in the proteasome lid and COP9 signalosome.

The large distance separating individual parts of eIF3 on the 40S subunit points to specialization of their functions. The main role of the PCI/MPN subunits seems to be scaffolding. It involves direct interaction with the 40S subunit, eIF1 (ref. 28) and eIF4G (refs 34, 35), and configuration of the non-core subunits within the same complex, while allowing them to interact with distant ribosomal regions. Interaction with eIF4G is essential and is thought to bridge the cap-binding complex, coupling its helicase activity with 43S com-

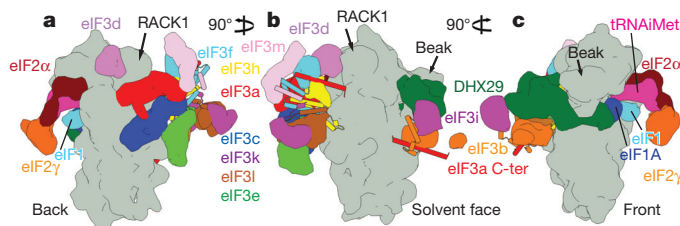


Figure 5 | Schematic representation of the arrangement of initiation factors and their subunits in the DHX29-bound 43S complex. a–c, Figure includes only eIFs and subunits for which the structures are known. The 40S subunit is depicted in grey surface; all other factors and subunits are labelled and coloured variably. eIF3 helical bundles fortifying the intersubunit interactions are represented as cylinders. The 43S complex is shown from the back (a), solvent side (b) and front (c).

plexes during attachment to mRNA and scanning. Initially, eIF4G was shown to interact with eIF3e (ref. 34), but recent analysis revealed a second, adjacent site on eIF4G that binds to eIF3c and eIF3d (ref. 35). This suggests that the binding site for eIF4G is located in a region of eIF3 where the c, d and e subunits come into proximity. Although in our density map eIF3d is separated from the PCI/MPN core, at least part of the unassigned density at the core (Extended Data Fig. 1g, red density) could belong to eIF3d. Importantly, regarding the position of eIF3e in 43S complexes, its interaction with eIF4G is also consistent with hydroxyl radical cleavage in ES6^S from the eIF4G middle domain³⁶. Taken together, the structure of the eIF3 core and biochemical data on the eIF3–eIF4G interaction allow us to suggest cautiously that eIF4F could reside on the solvent side of the 43S complex, perhaps between the PCI/MPN core and the eIF3b–eIF3i–eIF3g module.

Regarding the function of the eIF3b–eIF3i–eIF3g module, various studies implicated its subunits in scanning. Thus, eIF3b was suggested to be involved in start-codon selection^{37,38}, eIF3g in reinitiation by recycled 40S subunits³⁹, while eIF3i and eIF3g were suggested to stimulate scanning³⁹. The eIF3b–eIF3i–eIF3g module is situated at the mRNA channel entrance such that the mRNA would probably interact directly with module before entering its binding channel. Thus, these factors could form an extension of the entrance portion of the mRNA channel, facilitating mRNA entry and maintaining its position during scanning. eIF3b could also influence the conformation of h16, as it binds at its base. It is thus possible that one role of the eIF3b–eIF3i–eIF3g module is to maintain the mRNA entrance in a specific conformation. eIF3b also has a large interacting surface with DHX29. DHX29 most likely stimulates scanning indirectly, by inducing conformational changes in 43S complexes^{27,40}. Interaction with eIF3b could therefore establish proper ribosomal positioning of DHX29 and possibly also transmit to the 40S subunit conformational changes that are induced in DHX29 upon NTP hydrolysis.

As for eIF3d, in addition to interaction with eIF4G (ref. 35), its location near the mRNA exit and cross-linking to mRNA³¹ suggest that it might also participate in extending the exit portion of the mRNA channel.

In conclusion, our model provides a detailed view of the structure of mammalian eIF3 and of its interactions with components of the DHX29-bound 43S complex. It will serve as a framework for further elucidation of individual steps in eukaryotic translation initiation, which is crucial to a better understanding of translational control.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 8 March; accepted 6 July 2015.

Published online 7 September 2015.

- Jackson, R. J., Hellen, C. U. & Pestova, T. V. The mechanism of eukaryotic translation initiation and principles of its regulation. *Nature Rev. Mol. Cell Biol.* **11**, 113–127 (2010).
- Hinnebusch, A. G. eIF3: a versatile scaffold for translation initiation complexes. *Trends Biochem. Sci.* **31**, 553–562 (2006).
- Valášek, L. S. ‘Ribozoomin’-translation initiation from the perspective of the ribosome-bound eukaryotic initiation factors (eIFs). *Curr. Protein Pept. Sci.* **13**, 305–330 (2012).
- Ellisdon, A. M. & Stewart, M. Structural biology of the PCI-protein fold. *BioArchitecture* **2**, 118–123 (2012).
- Pena, V., Liu, S., Bujnicki, J. M., Lührmann, R. & Wahl, M. C. Structure of a multipartite protein-protein interaction domain in splicing factor prp8 and its link to retinitis pigmentosa. *Mol. Cell* **25**, 615–624 (2007).
- Enchev, R. I., Schreiber, A., Beuron, F. & Morris, E. P. Structural insights into the COP9 signalosome and its common architecture with the 26S proteasome lid and eIF3. *Structure* **18**, 518–527 (2010).
- Hashem, Y. *et al.* Structure of the mammalian ribosomal 43S preinitiation complex bound to the scanning factor DHX29. *Cell* **153**, 1108–1119 (2013).
- Hashem, Y. *et al.* Hepatitis-C-virus-like internal ribosome entry sites displace eIF3 to gain access to the 40S subunit. *Nature* **503**, 539–543 (2013).
- Querol-Audi, J. *et al.* Architecture of human translation initiation factor 3. *Structure* **21**, 920–928 (2013).
- Sun, C. *et al.* Functional reconstitution of human eukaryotic translation initiation factor 3 (eIF3). *Proc. Natl Acad. Sci. USA* **108**, 20473–20478 (2011).

11. Beck, F. *et al.* Near-atomic resolution structural model of the yeast 26S proteasome. *Proc. Natl Acad. Sci. USA* **109**, 14870–14875 (2012).
12. Lander, G. C. *et al.* Complete subunit architecture of the proteasome regulatory particle. *Nature* **482**, 186–191 (2012).
13. Lingaraju, G. M. *et al.* Crystal structure of the human COP9 signalosome. *Nature* **512**, 161–165 (2014).
14. ElAntak, L., Tzakos, A. G., Locker, N. & Lukavsky, P. J. Structure of eIF3b RNA recognition motif and its interaction with eIF3j: structural insights into the recruitment of eIF3b to the 40S ribosomal subunit. *J. Biol. Chem.* **282**, 8165–8174 (2007).
15. Herrmannová, A. *et al.* Structural analysis of an eIF3 subcomplex reveals conserved interactions required for a stable and proper translation pre-initiation complex assembly. *Nucleic Acids Res.* **40**, 2294–2311 (2012).
16. Liu, Y. *et al.* Translation initiation factor eIF3b contains a nine-bladed β -propeller and interacts with the 40S ribosomal subunit. *Structure* **22**, 923–930 (2014).
17. Erzberger, J. P. *et al.* Molecular architecture of the 40S-eIF1-eIF3 translation initiation complex. *Cell* **158**, 1123–1135 (2014).
18. Zhou, M. *et al.* Mass spectrometry reveals modularity and a complete subunit interaction map of the eukaryotic translation factor eIF3. *Proc. Natl Acad. Sci. USA* **105**, 18139–18144 (2008).
19. Dong, Z., Qi, J., Peng, H., Liu, J. & Zhang, J. T. Spectrin domain of eukaryotic initiation factor 3a is the docking site for formation of the a:b:i:g subcomplex. *J. Biol. Chem.* **288**, 27951–27959 (2013).
20. Wagner, S., Herrmannová, A., Malik, R., Peclínová, L. & Valášek, L. S. Functional and biochemical characterization of human eukaryotic translation initiation factor 3 in living cells. *Mol. Cell. Biol.* **34**, 3041–3052 (2014).
21. Masutani, M., Sonenberg, N., Yokoyama, S. & Imataka, H. Reconstitution reveals the functional core of mammalian eIF3. *EMBO J.* **26**, 3373–3383 (2007).
22. Fraser, C. S. *et al.* The j-subunit of human translation initiation factor eIF3 is required for the stable binding of eIF3 and its subcomplexes to 40S ribosomal subunits *in vitro*. *J. Biol. Chem.* **279**, 8946–8956 (2004).
23. Valášek, L., Hasek, J., Trachsel, H., Imre, E. M. & Ruis, H. The *Saccharomyces cerevisiae* HCR1 gene encoding a homologue of the p35 subunit of human translation initiation factor 3 (eIF3) is a high copy suppressor of a temperature-sensitive mutation in the Rpg1p subunit of yeast eIF3. *J. Biol. Chem.* **274**, 27567–27572 (1999).
24. Estrin, E., Lopez-Blanco, J. R., Chacon, P. & Martin, A. Formation of an intricate helical bundle dictates the assembly of the 26S proteasome lid. *Structure* **21**, 1624–1635 (2013).
25. Anger, A. M. *et al.* Structures of the human and *Drosophila* 80S ribosome. *Nature* **497**, 80–85 (2013).
26. Schmitt, E. *et al.* Structure of the ternary initiation complex aIF2-GDPNP-methionylated initiator tRNA. *Nature Struct. Mol. Biol.* **19**, 450–454 (2012).
27. Dhote, V., Sweeney, T. R., Kim, N., Hellen, C. U. & Pestova, T. V. Roles of individual domains in the function of DHX29, an essential factor required for translation of structured mammalian mRNAs. *Proc. Natl Acad. Sci. USA* **109**, E3150–E3159 (2012).
28. Aylett, C. H., Boehringer, D., Erzberger, J. P., Schaefer, T. & Ban, N. Structure of a yeast 40S-eIF1-eIF1A-eIF3-eIF3j initiation complex. *Nature Struct. Mol. Biol.* **22**, 269–271 (2015).
29. Kouba, T. *et al.* Small ribosomal protein RPS0 stimulates translation initiation by mediating 40S-binding of eIF3 via its direct contact with the eIF3a/TIF32 subunit. *PLoS ONE* **7**, e40464 (2012).
30. Valášek, L. *et al.* The yeast eIF3 subunits TIF32/a, NIP1/c, and eIF5 make critical connections with the 40S ribosome *in vivo*. *Genes Dev.* **17**, 786–799 (2003).
31. Pisarev, A. V., Kolupaeva, V. G., Yusupov, M. M., Hellen, C. U. & Pestova, T. V. Ribosomal position and contacts of mRNA in eukaryotic translation initiation complexes. *EMBO J.* **27**, 1609–1621 (2008).
32. Zhou, C. *et al.* PCI proteins eIF3e and eIF3m define distinct translation initiation factor 3 complexes. *BMC Biol.* **3**, 14 (2005).
33. Rezende, A. M. *et al.* The translation initiation complex eIF3 in trypanosomatids and other pathogenic excavates-identification of conserved and divergent features based on orthologue analysis. *BMC Genomics* **15**, 1175 (2014).
34. LeFebvre, A. K. *et al.* Translation initiation factor eIF4G-1 binds to eIF3 through the eIF3E subunit. *J. Biol. Chem.* **281**, 22917–22932 (2006).
35. Villa, N., Do, A., Hershey, J. W. B. & Fraser, C. S. Human eukaryotic initiation factor 4G (eIF4G) binds to eIF3c, -d, and -e to promote mRNA recruitment to the ribosome. *J. Biol. Chem.* **288**, 32932–32940 (2013).
36. Yu, Y., Abaeva, I. S., Marintchev, A., Pestova, T. V. & Hellen, C. U. Common conformational changes induced in type 2 picornavirus IRESs by cognate trans-acting factors. *Nucleic Acids Res.* **39**, 4851–4865 (2011).
37. Chiu, W. L. *et al.* The C-terminal region of eukaryotic translation initiation factor 3a (eIF3a) promotes mRNA recruitment, scanning, and, together with eIF3j and the eIF3b RNA recognition motif, selection of AUG start codons. *Mol. Cell. Biol.* **30**, 4415–4434 (2010).
38. Elantak, L. *et al.* The indispensable N-terminal half of eIF3j/HCR1 cooperates with its structurally conserved binding partner eIF3b/PRT1-RRM and with eIF1A in stringent AUG selection. *J. Mol. Biol.* **396**, 1097–1116 (2010).
39. Cuchalová, L. *et al.* The RNA recognition motif of eukaryotic translation initiation factor 3g (eIF3g) is required for resumption of scanning of posttermination ribosomes for reinitiation on GCN4 and together with eIF3i stimulates linear scanning. *Mol. Cell. Biol.* **30**, 4671–4686 (2010).
40. Pisareva, V. P., Pisarev, A. V., Komar, A. A., Hellen, C. U. & Pestova, T. V. Translation initiation on mammalian mRNAs with structured 5'UTRs requires DExH-box protein DHX29. *Cell* **135**, 1237–1250 (2008).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank R. A. Grassucci for helping collect the electron microscopy images, H. Kao for help with computations, and the proteomic platform Strasbourg-Esplanade for conducting the nano-LS-MS/MS analysis. This work was supported by both HHMI and NIH R01 GM29169 (to J.F.) and NIH R01 GM59660 (to T.V.P.). This work has been published under the framework of the LABEX: ANR-10-LABX-0036_NETRNA and benefits from funding from the state managed by the French National Research Agency as part of the Investments for the future program (to Y.H.).

Author Contributions A.d.G., V.D., L.K., C.U.T.H., T.V.P., J.F. and Y.H. interpreted the data and wrote the manuscript. V.D. and T.V.P. prepared the samples. Y.H. and A.d.G. performed the cryo-electron microscopy experiments, data processing and the atomic modelling. L.K. performed the mass-spectrometry experiment. T.V.P., J.F. and Y.H. directed research.

Author Information Cryo-EM maps and atomic coordinates for the reported structures have been deposited in the Electron Microscopy Data Bank under the accession numbers EMD-3056, EMD-3057 and EMD-3058, and in the Protein Data Bank under the accession numbers 5A5T and 5A5U. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to T.V.P. (Tatyana.Pestova@downstate.edu), J.F. (jf2192@cumc.columbia.edu) or Y.H. (y.hashem@ibmc-cnrs.unistra.fr).

METHODS

No statistical methods were used to predetermine sample size.

Purification of 40S ribosomal subunits, initiation factors, DHX29 and aminoacylation of tRNA_i^{Met}. The exact protocols have been described previously⁷.

Assembly and purification of 43S–DHX29 preinitiation complexes. The exact protocols have been described previously⁷.

Electron microscopy. Three microlitres of each sample was applied to holey carbon grids (carbon-coated Quantifoil 2/4 grid, Quantifoil Micro Tools GmbH) containing an additional continuous thin layer of carbon⁴¹ without previous plasma cleaning. Grids were blotted and vitrified by rapidly plunging into liquid ethane at -180°C with a Vitrobot Mark IV (FEI)⁴². Data acquisition was done under low-dose conditions ($25\text{ e}^{-}\text{Å}^{-2}$)⁴³ on a FEI Tecnai F30 Polara (FEI) operating at 300 kV. The data set was collected with the automated data collection system Leginon⁴⁴ at a calibrated magnification of $\times 30,120$ on a Gatan K2 Summit direct electron detection camera with a pixel size of 1.66 Å on the object scale. A total of 11,274 micrographs were recorded in electron-counting, dose-fractionation mode as a series of 20 frames, with an accumulation time of 0.4 s per frame. The dose rate was set to 8 counts per physical pixel per s ($\sim 10\text{ e}^{-}\text{pixel}^{-1}\text{ s}^{-1}$). The total exposure time was 8 s, leading to a total accumulated dose of approximately $25\text{ e}^{-}\text{Å}^{-2}$ on the specimen.

Image processing. Dose-fractionated image stacks were aligned using dosegpu_driftcorr⁴⁵ and the sum of frames 3–20 in each image stack was used for further processing. The first two frames were omitted as they contained motions greater than 4 Å per frame on average. The data were then preprocessed using pySPIDER and Arachnid (<http://www.arachnid.us>)⁷. Arachnid is a Python-encapsulated version of SPIDER⁴⁶ replacing SPIDER batch files with Python scripts. It also contains novel procedures such as Autopicker⁴⁷, which was used for fully automated particle selection. Approximately 652,000 particles were extracted from the data set. These particles were classified with RELION⁴⁸, yielding two high-quality classes containing 272,252 particles, or $\sim 42\%$ of the data set. The remaining particles were rejected (Extended Data Fig. 2). One class (19% of the particles) displays a 40S ribosomal subunit with eIF3 and the eIF2–Met-tRNA_i^{Met} ternary complex bound, while the other (23% of the particles) displays a bare 40S ribosomal subunit. The class containing eIF3 was further classified, and all classes showing eIF3 with a density of good quality were pooled together (84,850 particles) and refined to a resolution of 6.2 Å (Extended Data Fig. 1b). Reported resolutions were calculated with the FSC = 0.143 criterion⁴⁹ using the ‘gold standard’ protocol, ensuring independence of half-set reconstructions. The reported resolution measurements were corrected for the effects of a soft mask on the FSC curve using high-resolution noise substitution⁵⁰. Before visualization, all density maps were sharpened by applying a negative B-factor that was estimated using automated procedures⁴⁹. The local resolution, as measured using the program ResMap⁵¹, varies between 5 and 12 Å across most of the map (Extended Data Fig. 2d, f). Particles were then realigned to the eIF3 core as a reference using a soft mask around eIF3 core subunits. This procedure improved the resolution of the more distal regions of eIF3, and resulted in a map of eIF3 core subunits at local resolutions estimated by ResMap to be between 6 Å for most of eIF3 core and up to 8 Å (Extended Data Fig. 2c, e). These two maps were used for modelling the eIF3 core subunits. Next, focused classification was performed to isolate eIF3 peripheral subunits in two locations: in the vicinity of DHX29 and at the back of the ribosome head, close to RACK1. In the focused classification, we used masks with smooth edges created by UCSF Chimera⁵² and SEGGER⁵³, and encompassing large regions around the regions of interest (Extended Data Fig. 8). To avoid diverging orientation assignments during the Relion classification, owing to the small volume encompassed by the mask compared to the total volume, particle orientation assignments were kept fixed. The classification with a mask in the region of DHX29 gave one major class (51% of the particles) with well-defined eIF3 peripheral subunits in the vicinity of DHX29 (Extended Data Fig. 8a), which was subsequently refined to a resolution of 7.1 Å (Extended Data Fig. 1b). The classification with a mask around the back of the 40S head gave one class (18% of the particles) with a strong density in this region, in contact with RACK1 (Extended Data Fig. 8b). Further classification of particles in this class reduced the conformational variability of the density and isolated one major conformer (39% of the remaining particles). This class of particles, a total of 10,062 particles, yielded a reconstruction at 7.7 Å (Extended Data Fig. 1b).

Identification and annotation of eIF3 domains. The different domains of the eIF3 constituent subunits were annotated (Extended Data Fig. 1a) on the basis of our eIF3 polyaniline-level model where possible. Thus, the coordinates of the WD40, RRM and PCI domains were annotated based on our model. Alternatively, eIF3 domains were detected mainly using the conserved domain detector (CDD) tool⁵⁴, which produces a graphical display of conserved domains identified in the protein using RPS-BLAST, and the InterPro protein sequence analysis and classification online tool⁵⁵. The annotation relies on the top-scoring

hits by default and represents a good approximation of the domain coordinates. The sequences of rabbit eIF3 subunits used for domain detection had the following NCBI accession numbers: gi|655868267 for eIF3a, gi|291390868 for eIF3c, gi|655763812 for eIF3d, gi|291388413 for eIF3e, gi|655602713 for eIF3f, gi|655743367 for eIF3g, gi|291388434 for eIF3h, gi|75070231 for eIF3i, gi|655862979 for eIF3j, gi|291390078 for eIF3k, gi|291414663 for eIF3l and gi|291384781 for eIF3m. Note that the human eIF3b sequence (gi|83367072) was used rather than the rabbit sequence, because our nanoscale liquid chromatography coupled to tandem mass spectrometry (nano-LC–MS/MS) analysis (Extended Data Fig. 9c) raised doubts about the length of the N-terminal region of the latter shown in currently available database entries. In addition to domains identified using CDD (Extended Data Fig. 1a), a spectrin domain was added to eIF3a (ref. 19) and a zinc-binding domain was added to eIF3g (ref. 56).

In-gel protein digestion for nanoLC–MS/MS analysis. The protocol has been described previously⁵⁷. In brief, SDS–PAGE bands were excised from the gel and transferred into 96-well microtitre plates. The bands were then destained and dehydrated. Gel pieces were washed again with the destaining solutions and then incubated with trypsin (Promega) and chymotrypsin (Promega) for digestion overnight at room temperature. The resulting peptides were extracted from the gel pieces. The initial digestion and extraction supernatants were pooled together and vacuum-dried in a SpeedVac concentrator.

Nano-LC–electrospray-ionization triple-TOF MS/MS analysis. Dried tryptic digests were resuspended in 12 µl of water containing 0.1% formic acid (solvent A) before analysis on a NanoLC-2DPlus system (with nanoFlex Chip module; Eksigent, Sciex Separations) coupled to a TripleTOF 5600 mass spectrometer (AB Sciex) operating in positive mode. Five microlitres of each sample was loaded on a Chip C-18 precolumn ($300\text{ µm ID} \times 5\text{ mm ChromXP}$; Eksigent) at 2 µl min^{-1} in solvent A. After 10 min of desalting and concentration in the trap, the system was switched online with the analytical Chip C-18 analytical column ($75\text{ µm ID} \times 15\text{ cm ChromXP}$; Eksigent) equilibrated in 95% solvent A and 5% solvent B (0.1% formic acid in acetonitrile). Peptides were eluted by using a 5–40% gradient of solvent B for 60 min at a flow rate of 300 nl min^{-1} . The TripleTOF 5600 was operated in data-dependent acquisition mode with Analyst software (version 1.5, ABSciex). Survey MS scans were acquired during 250 ms in the 400–1,250 m/z range. Up to 20 of the most intense multiply charged ions ($2+$ to $5+$) were selected for CID (collision-induced dissociation) fragmentation, if they exceeded the 150 counts per second intensity threshold. Ions were fragmented using a rolling collision energy within a 60 ms accumulation time and an exclusion time of 15 s. This ‘top20’ method, with a constant cycle time of 1.5 s, was set in high-sensitivity mode.

To obtain optimal mass accuracy, a β -galactosidase digest (AB Sciex) was injected before each sample using the ‘Autocal’ feature from Analyst software: calibration was performed using the 10 more abundant peptides in MS mode and with the 729.3652 m/z precursor in MS/MS mode. Moreover, to prevent carry-over due to stationary phase memory, two consecutive washing runs were performed after each sample injection, as well as a blank injection (solvent A) to verify that no peptides were identified due to a carry-over phenomenon.

Database search and data analysis. The protocol has been described previously⁵⁷. Data were searched against the complete human and rabbit proteomes set from the SwissProt database. The algorithm used for database search was Mascot⁵⁸ (version 2.2, Matrix Science) through the ProteinScape package⁵⁹ (v3.1, Bruker). Peptide modifications allowed during the search were: N-acetyl (protein), carbamidomethylation (C) and oxidation (M). Mass tolerances in MS and MS/MS were set to 30 p.p.m. and 0.5 Da, respectively, and the instrument setting was specified as ESI-QUAD-TOF. Three trypsin or chymotrypsin missed cleavages were allowed. A decoy database strategy⁶⁰ was used to validate Mascot identifications at FDR $< 1\%$ (individual identity scores varied between 30 and 34 for each data search) using the ProteinScape Protein Assessment tool. Data were further validated by manually inspecting the quality of the MS/MS fragmentation spectra: a minimum of five consecutive amino acids was requested to validate the spectrum, as well as other rules (proline-specific fragmentation pattern, major peaks assigned to fragments, and so on). To compare the position of the validated peptides on the full-length sequences of the different proteins, the Protein Sequence viewer from ProteinScape package was used. Moreover, to obtain maximal sequence coverage on the identified proteins, a second round of database search was set up without defining enzyme specificity: semi-trypsin or non-trypsin peptides were thus highlighted in the trypsin digest analysis, as for the SDS–PAGE bands digested by chymotrypsin.

Segmentation and display of density maps. Cryo-EM reconstructions were segmented using the SEGGER module⁵³ implemented in UCSF Chimera⁵². Segments counting less than 10,000 voxels were discarded. Segments were refined manually using the VOLUME ERASER module implemented in UCSF Chimera. Finally, the segments obtained were smoothed using a Gaussian filter in the VOLUME FILTER module also implemented in Chimera. The final maps were displayed and rendered with Chimera.

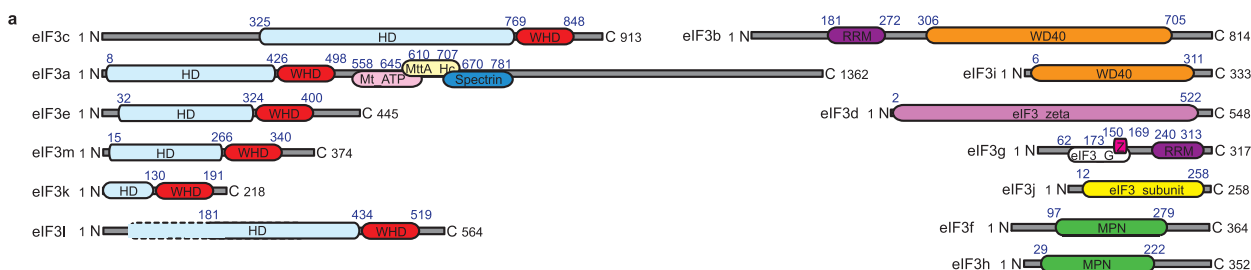
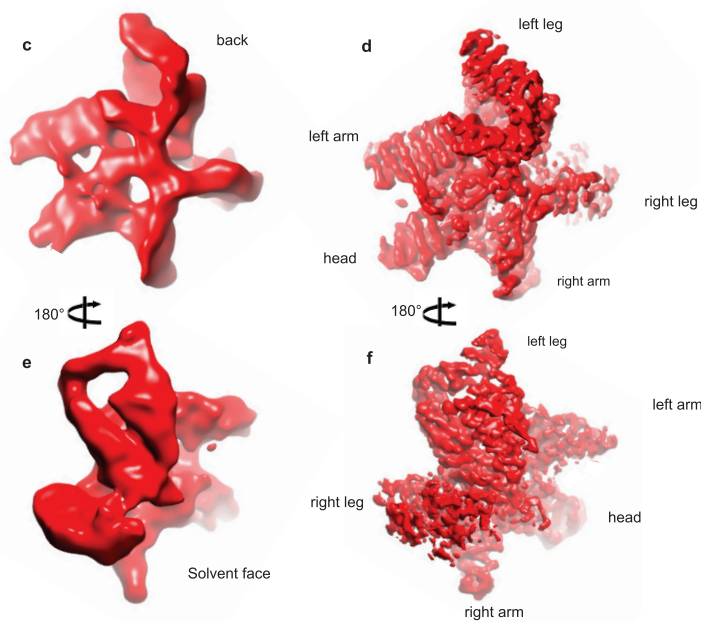
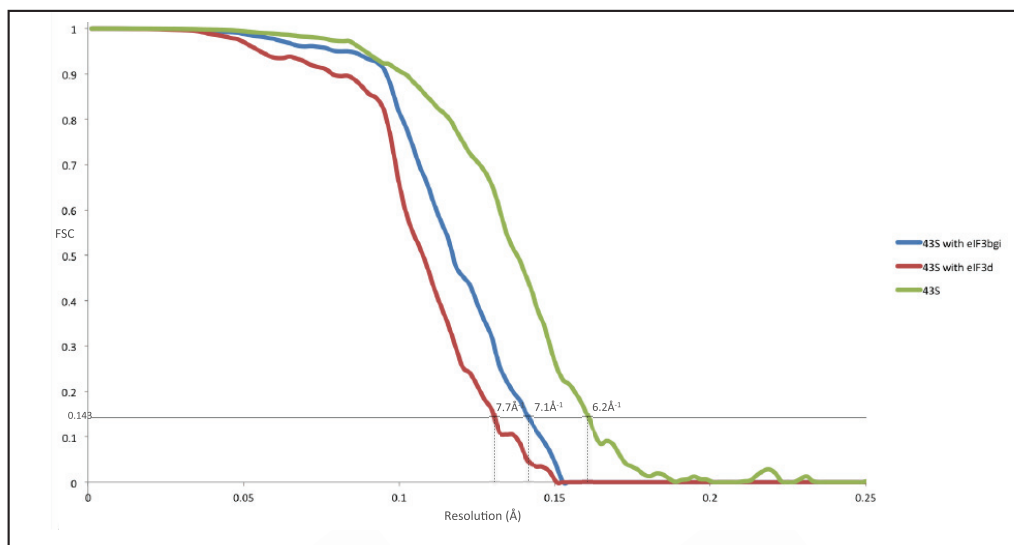
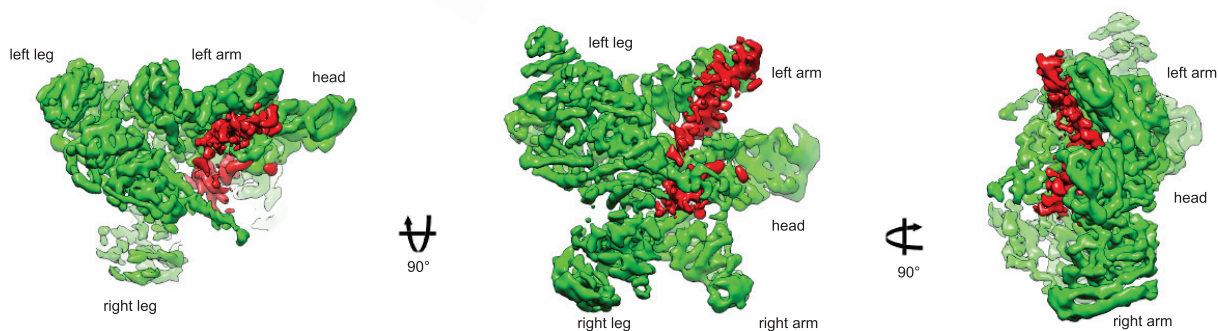
Atomic modelling of eIF3. We were able to model residues 7–605 of eIF3a (~43%), 320–876 of eIF3c (~61%), 4–422 of eIF3e (~94%), 93–364 of eIF3f (~75%), 29–352 of eIF3h (~92%), 2–216 of eIF3k (~98%), 181–552 of eIF3l (~66%) and 7–370 of eIF3m (~97%). To generate an atomic model of eIF3, each subunit was separately modelled into its corresponding density segment. We should stress that our cryo-EM map does not allow atomic modelling of side-chains and exact residue registration, and therefore our model of eIF3 should be considered as a polyaniline-level model. eIF3a was modelled by homology to crystal structures of *S. cerevisiae* eIF3a (refs 17, 61). The additional long helix in the C-terminal tail was modelled *ab initio* by first predicting its secondary structure using the SYMPRED web service tool (<http://www.ibi.vu.nl/programs/sympredwww/>), and then generating the predicted long helix in SWISS-MODEL⁶². The 3D structure generated was then fitted into its cryo-EM density in UCSF Chimera³² by manually modifying the backbone torsion angles using the ADJUST TORSION module implemented in Chimera. The structure of the whole subunit was then refitted using molecular dynamics flexible fitting as described in the next section.

Similarly, eIF3c was modelled by homology to the crystal structure of its *S. cerevisiae* counterpart¹⁷. The variant parts of its structure along with the helix in the C terminus tail were modelled as described above for eIF3a. Rabbit eIF3k was modelled entirely by homology to the crystal structure of human eIF3k⁶³. The X-ray structures of the human COP9 subunits CSN1, CSN5 and CSN6 (ref. 13) were used as templates for modelling rabbit eIF3e, eIF3h and eIF3f; the variant parts of the latter subunits, compared to their modelling templates, were modelled according to the method described above for eIF3a, regarding the modelling of the C-terminal helical tails. Modelling of the remaining eIF3m and eIF3l subunits was based mainly on the atomic model of the proteasome lid subunits RPN9 and RPN3 (ref. 11) owing to their closely related topology, as indicated by our cryo-EM reconstruction, and variant parts were modelled *ab initio* according to the method described for eIF3a. For most parts that were modelled *ab initio*, the secondary structure elements predicted and validated by the cryo-EM reconstruction map were helical.

As for eIF3 peripheral subunits, based on our density map, the crystal structures of the eIF3b WD40 domain from *Chaetomium thermophilum*¹⁶, the C-terminal helical domain of *S. cerevisiae* eIF3b in complex with eIF3i (seven-bladed β -propeller)¹⁵ and the solution NMR structure of the RRM domain of the N-terminal end of human eIF3b (ref. 14), were used to derive a homology model of nearly the full eIF3b from rabbit (residues 73–685, ~85%, using the rabbit eIF3b sequence¹⁶, gi|291415469). This fragment corresponds to residues 170–782 of human eIF3b; for further discussion of the eIF3b sequence see Extended Data Fig. 9c. Yeast eIF3i crystal structure¹⁵ was rigid-body fitted using the FIT IN MAP module implemented in Chimera³². We did not attempt to generate a homology model of rabbit eIF3i due to the low local resolution of the cryo-EM reconstruction in that region and simply fitted the crystal structure of yeast eIF3i as a rigid body.

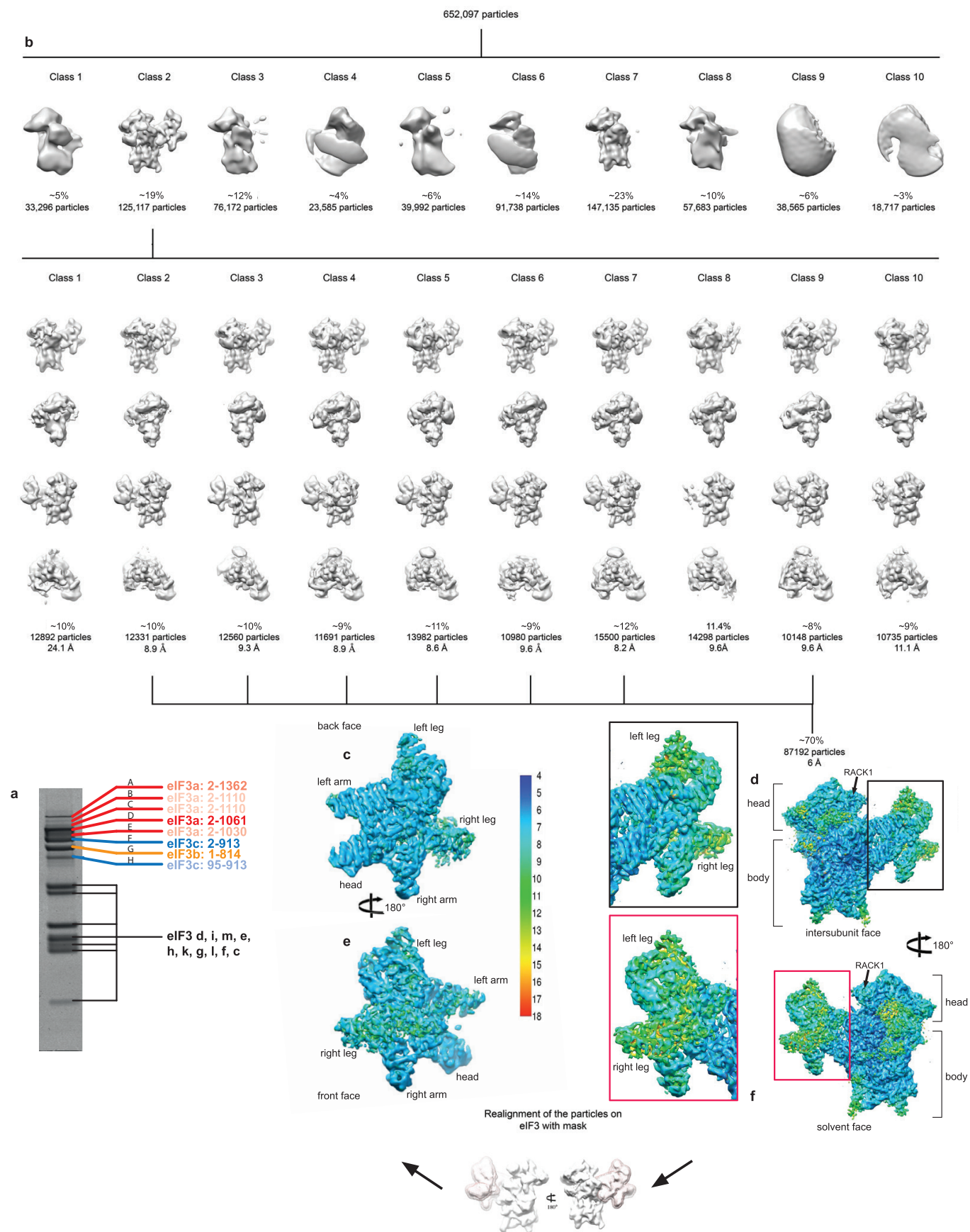
Fitting of atomic structures into electron microscopy maps. After the construction of the eIF3 polyaniline-level model, it was placed into its cryo-EM density map along with the atomic model of the 40S subunit of the human ribosome²⁵ by rigid-body fitting using Chimera as described above. Remaining parts of the 43S cryo-EM density map corresponding to the eIF2-ternary complex, eIF3b, eIF3i and DHX29 were simply segmented out using the VOLUME ERASER module implemented in UCSF Chimera. Starting from this system, everything was flexibly fitted into the map using MDFF (molecular dynamics flexible fitting)⁶⁴ as described in previous work⁶⁵. In brief, the initial system was prepared for MDFF using VMD⁶⁶ and the trajectories run in NAMD^{66,67}. To achieve a better representation of the inter- and intra-molecular interactions, the system was embedded in a solvent box of TIP3P water molecules, with an extra 12 Å padding in each direction, and neutralized by potassium ions, and an excess of ~0.2 M KCl was added. The simulated system was prepared using CHARMM force field parameters (Combined CHARMM All-Hydrogen Topology File for CHARMM22 Proteins and CHARMM27 Lipids)^{67,68}. The trajectories were run in explicit solvent. The run was stopped at 600 ps of simulation time.

41. Grassucci, R. A., Taylor, D. J. & Frank, J. Preparation of macromolecular complexes for cryo-electron microscopy. *Nature Protocols* **2**, 3239–3246 (2007).
42. Dubochet, J. *et al.* Cryo-electron microscopy of vitrified specimens. *Q. Rev. Biophys.* **21**, 129–228 (1988).
43. Wagenknecht, T., Frank, J., Boublik, M., Nurse, K. & Ofengand, J. Direct localization of the tRNA-anticodon interaction site on the *Escherichia coli* 30 S ribosomal subunit by electron microscopy and computerized image averaging. *J. Mol. Biol.* **203**, 753–760 (1988).
44. Suloway, C. *et al.* Automated molecular microscopy: the new Leginon system. *J. Struct. Biol.* **151**, 41–60 (2005).
45. Li, X. *et al.* Electron counting and beam-induced motion correction enable near-atomic-resolution single-particle cryo-EM. *Nature Methods* **10**, 584–590 (2013).
46. Frank, J. *et al.* SPIDER and WEB: processing and visualization of images in 3D electron microscopy and related fields. *J. Struct. Biol.* **116**, 190–199 (1996).
47. Langlois, R. *et al.* Automated particle picking for low-contrast macromolecules in cryo-electron microscopy. *J. Struct. Biol.* **186**, 1–7 (2014).
48. Scheres, S. H. RELION: implementation of a Bayesian approach to cryo-EM structure determination. *J. Struct. Biol.* **180**, 519–530 (2012).
49. Rosenthal, P. B. & Henderson, R. Optimal determination of particle orientation, absolute hand, and contrast loss in single-particle electron cryomicroscopy. *J. Mol. Biol.* **333**, 721–745 (2003).
50. Chen, S. *et al.* High-resolution noise substitution to measure overfitting and validate resolution in 3D structure determination by single particle electron cryomicroscopy. *Ultramicroscopy* **135**, 24–35 (2013).
51. Kucukelbir, A., Sigworth, S. J. & Tagare, D. Quantifying the local resolution of cryo-EM density maps. *Nature Methods* **11**, 63–65 (2014).
52. Pettersen, E. F. *et al.* UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).
53. Pintilie, G. D., Zhang, J., Goddard, T. D., Chiu, W. & Gossard, D. C. Quantitative analysis of cryo-EM density map segmentation by watershed and scale-space filtering, and fitting of structures by alignment to regions. *J. Struct. Biol.* **170**, 427–438 (2010).
54. Marchler-Bauer, A. *et al.* CDD: NCBI's conserved domain database. *Nucleic Acids Res.* **43**, D222–D226 (2015).
55. Mitchell, A. *et al.* The InterPro protein families database: the classification resource after 15 years. *Nucleic Acids Res.* **43**, D213–D221 (2015).
56. Marintchev, A. & Wagner, G. Translation initiation: structures, mechanisms and evolution. *Q. Rev. Biophys.* **37**, 197–284 (2004).
57. Lange, H. *et al.* The RNA helicases AtMTR4 and HEN2 target specific subsets of nuclear transcripts for degradation by the nuclear exosome in *Arabidopsis thaliana*. *PLoS Genet.* **10**, e1004564 (2014).
58. Perkins, D. N., Pappin, D. J., Creasy, D. M. & Cottrell, J. S. Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis* **20**, 3551–3567 (1999).
59. Thiele, H., Glandorf, J. & Hufnagel, P. Bioinformatics strategies in life sciences: from data processing and data warehousing to biological knowledge extraction. *J. Integr. Bioinform.* **7**, 141 (2010).
60. Elias, J. E. & Gygi, S. P. Target-decoy search strategy for mass spectrometry-based proteomics. *Methods Mol. Biol.* **604**, 55–71 (2010).
61. Khoshnevis, S. *et al.* Structural integrity of the PCI domain of eIF3a/TIF32 is required for mRNA recruitment to the 43S pre-initiation complexes. *Nucleic Acids Res.* **42**, 4123–4139 (2014).
62. Guex, N. & Peitsch, M. C. SWISS-MODEL and the Swiss-PdbViewer: An environment for comparative protein modeling. *Electrophoresis* **18**, 2714–2723 (1997).
63. Wei, Z. *et al.* Crystal structure of human eIF3k, the first structure of eIF3 subunits. *J. Biol. Chem.* **279**, 34983–34990 (2004).
64. Trabuco, L. G., Villa, E., Mitra, K., Frank, J. & Schulten, K. Flexible fitting of atomic structures into electron microscopy maps using molecular dynamics. *Structure* **16**, 673–683 (2008).
65. Hashem, Y. *et al.* High-resolution cryo-electron microscopy structure of the *Trypanosoma brucei* ribosome. *Nature* **494**, 385–389 (2013).
66. Humphrey, W., Dalke, A. & Schulten, K. VMD: visual molecular dynamics. *J. Mol. Graph.* **1**, 33–8 27–8 (1996).
67. Brooks, B. R. *et al.* CHARMM: a program for macromolecular energy, minimization, and dynamics calculations. *J. Comput. Chem.* **4**, 187–217 (1983).
68. MacKerell, A. D. Jr. *et al.* in *The Encyclopedia of Computational Chemistry* (eds von Rague Schleyer, P. *et al.*) 271–277 (John Wiley & Sons, 1998).

**b****g**

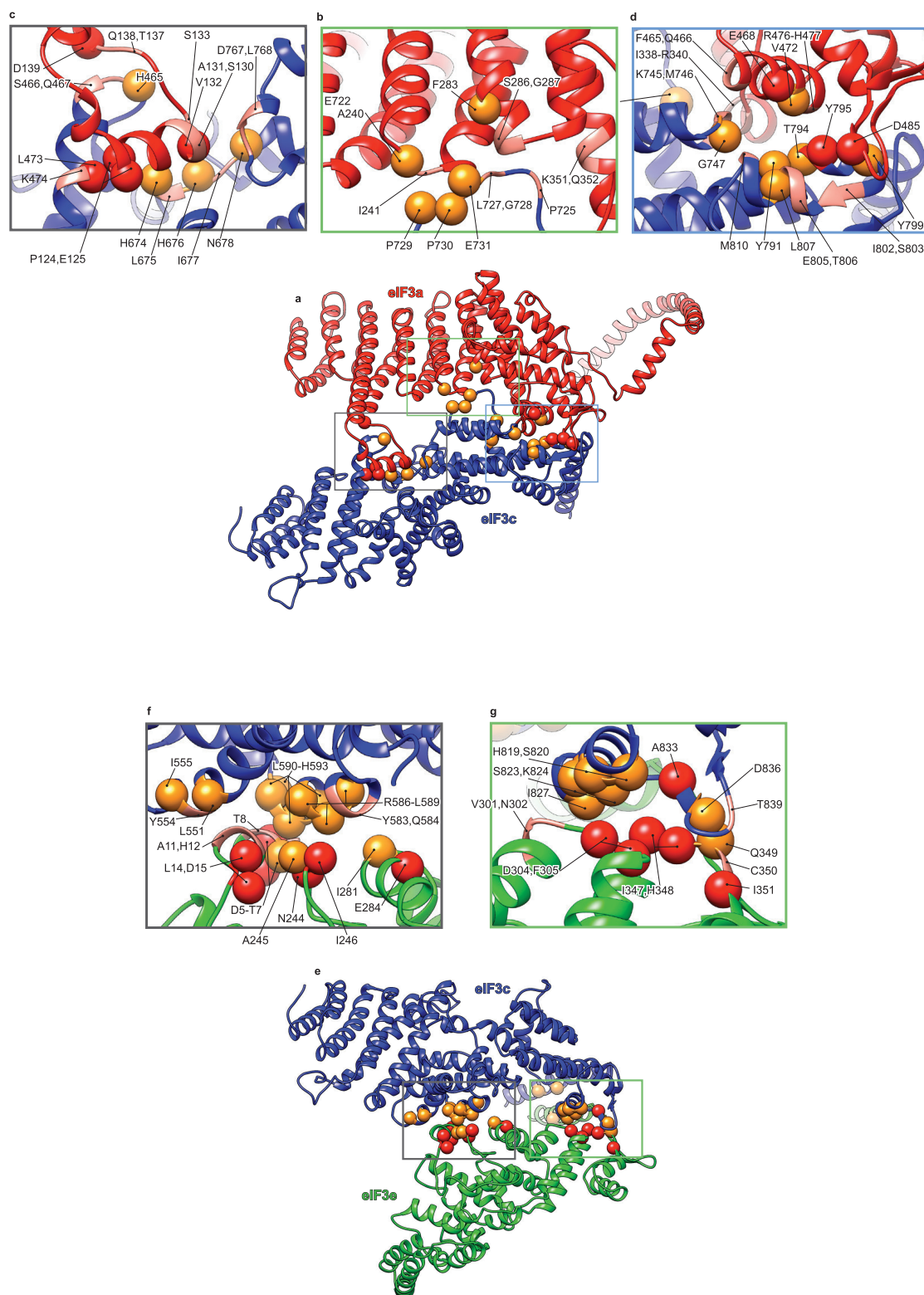
Extended Data Figure 1 | Domain organization of the eIF3 subunits and resolution of eIF3 core. **a**, Schematic representation of the domain organization of the rabbit eIF3 subunits (see Methods). Domain boundaries are indicated, and based where possible on our polyaniline-level model of eIF3. HD, helical domain; Z, zinc-recognition motif. Dashed line in eIF3l subunit diagram indicates that the helical domain might extend further in the N terminus, but it was not possible to be conclusive based on our density map. **b**, Fourier shell correlation (FSC) of the different 43S complex reconstructions used during analysis. The resolution estimation followed the 'gold standard' protocol ensuring independence of the half-set reconstructions. *x* axis, resolution in Å; *y* axis, FSC. Green line denotes 43S complexes including particles that present the eIF3 core subunits; blue line denotes 43S complexes

including all particles presenting DHX29 and the eIF3 peripheral subunits b, g and i; and red line denotes 43S complexes including all particles presenting the density attributed to eIF3d. For each reconstruction, a dashed line marks the resolution at which the FSC reaches the value of 0.143. **c–f**, Qualitative comparison of eIF3 core resolution in the present and in previous structures. **c, d**, eIF3 core cryo-EM structure from our previous study at 11.6 Å (ref. 7). **e, f**, eIF3 core cryo-EM structure from the present study at 6 Å after focused refinement. The eIF3 core is labelled according to the anthropomorphic nomenclature. **g**, Unassigned density region of the eIF3 core cryo-EM structure, coloured in red, seen from three different orientations. The green surface represents most of the core region that was modelled.



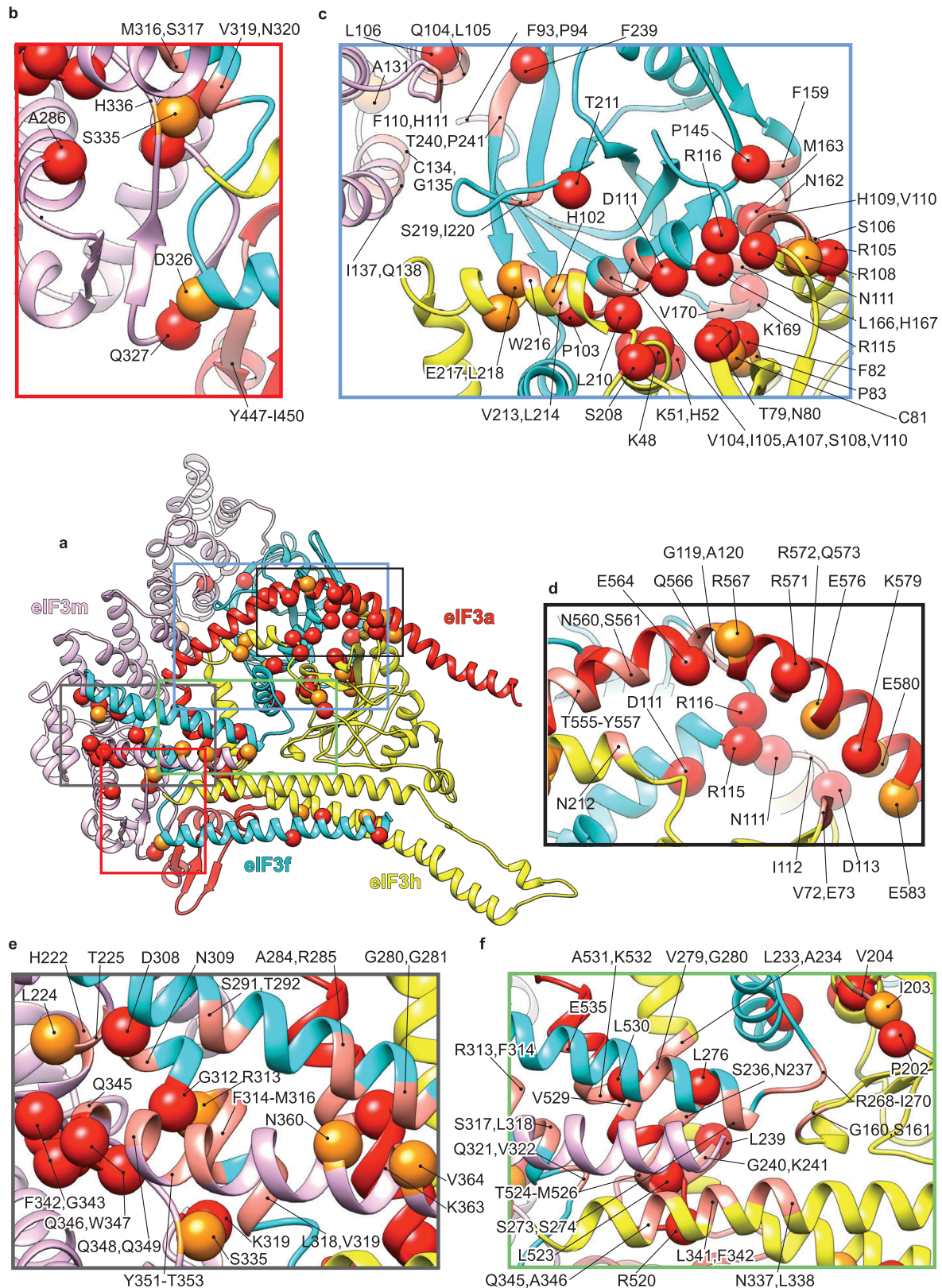
Extended Data Figure 2 | Sorting of particle images and focused refinement of the five-lobed core of eIF3. **a**, Composition of eIF3 purified from RRL for assembly of 43S complexes, resolved by SDS-PAGE and analysed by nano-LC-MS/MS to characterize truncation of eIF3a and eIF3c due to endoproteolytic cleavage. The intensity of labels in this panel reflects the intensity of bands corresponding to the truncated forms of eIF3a and eIF3c. The sequence of the N-terminal region of rabbit eIF3b has not been conclusively established (see Extended Data Fig. 9c) and numbering therefore refers to human eIF3b. **b**, Overview of the process of sorting particle images. The population of each class is indicated by the number of particles and the percentage of the total number of particles at the beginning of each of the two classification rounds. The DHX29-bound 43S complex was processed from a total of ~650,000

particle images, which were first sorted into ten different classes. Class 2 (~125,000 particles) was sorted into ten subclasses, which are displayed in four different orientations, showing the intersubunit face, front, solvent side and bottom, respectively. Classes 2–7 and 9 from the second classification round were pooled and refined yielding a reconstruction presenting a variable resolution ranging from 4.5 to 15 Å (bottom right). **d**, **f**, Cryo-EM reconstruction of the 43S complex, coloured according to the local resolution as measured using ResMap (see Methods). The red and black boxes correspond to close-up views of the eIF3 core viewed from the intersubunit face (**d**) and the solvent face (**f**) of the 40S subunit. **c**, **e**, CryoEM reconstruction of the eIF3 core after focused refinement, coloured according to the local resolution as measured using ResMap. Maps in **c–f** are filtered to 6 Å.



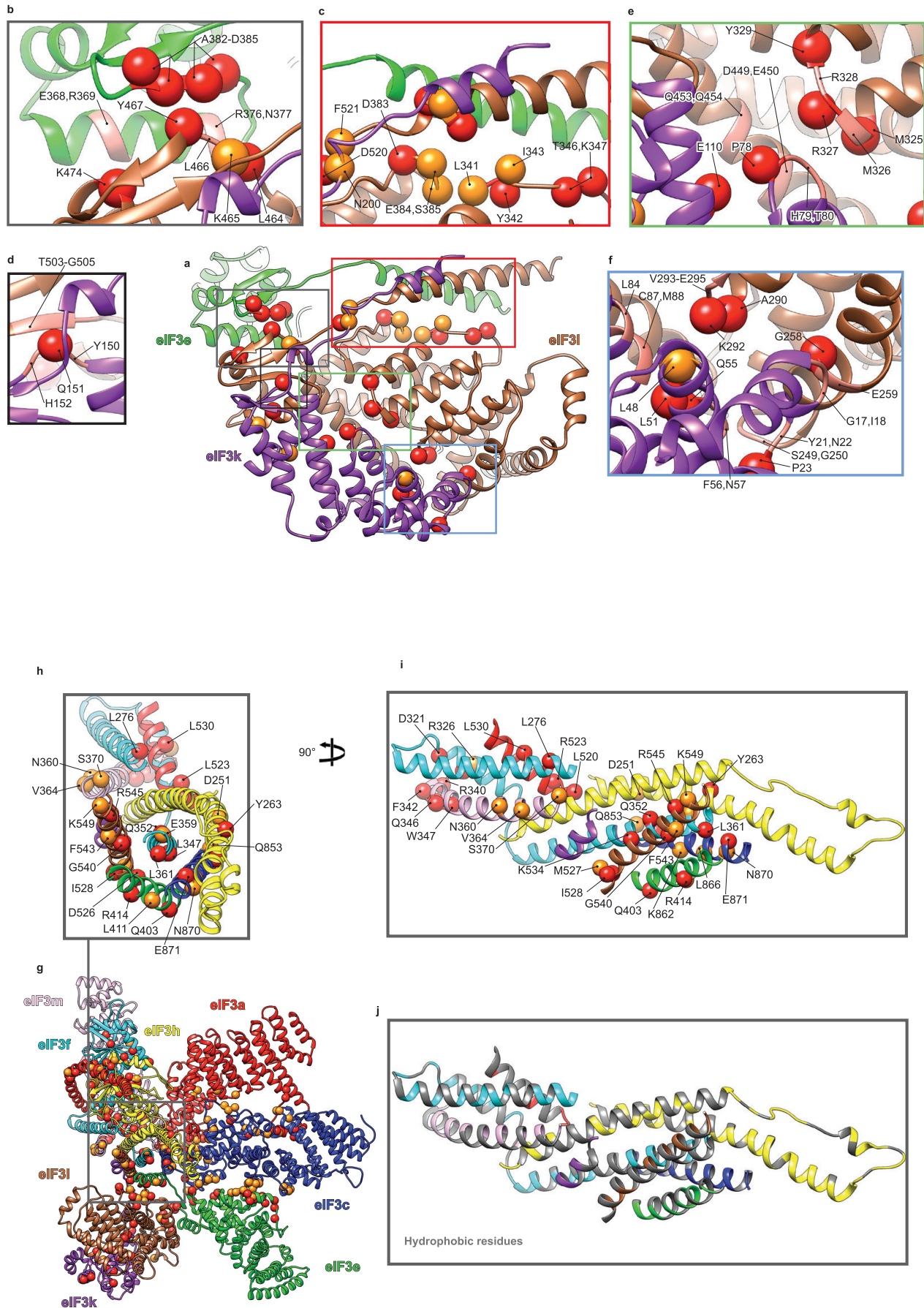
Extended Data Figure 4 | Conservation of quaternary interactions between eIF3a and eIF3c subunits, and between eIF3c and eIF3e subunits. **a**, Ribbon representation of eIF3a and eIF3c subunits. **b–d**, Close-up views of contact regions between eIF3a and eIF3c. **e**, Ribbon representation of eIF3c and eIF3e subunits. **f, g**, Close-up views of contact regions between subunits c and e of eIF3. Red spheres represent residues at the interfaces that are conserved in eIF3 from six representative eukaryotic organisms; *H. sapiens*, *C. elegans*, *A. thaliana*, *D. melanogaster* and *X. tropicalis*, which are very different

multicellular eukaryotic organisms, and *N. crassa*, a unicellular organism. These organisms all have a full complement of 13 eIF3 subunits. Orange spheres represent residues at the interfaces that are conserved only in the five multicellular eukaryotic organisms. The remaining residues that are suggested by the model and the density map to be involved in quaternary interactions are represented as ribbons in salmon colour. Many of these other residues are conserved in three or four of the compared organisms, but almost all of them have conserved chemical properties.



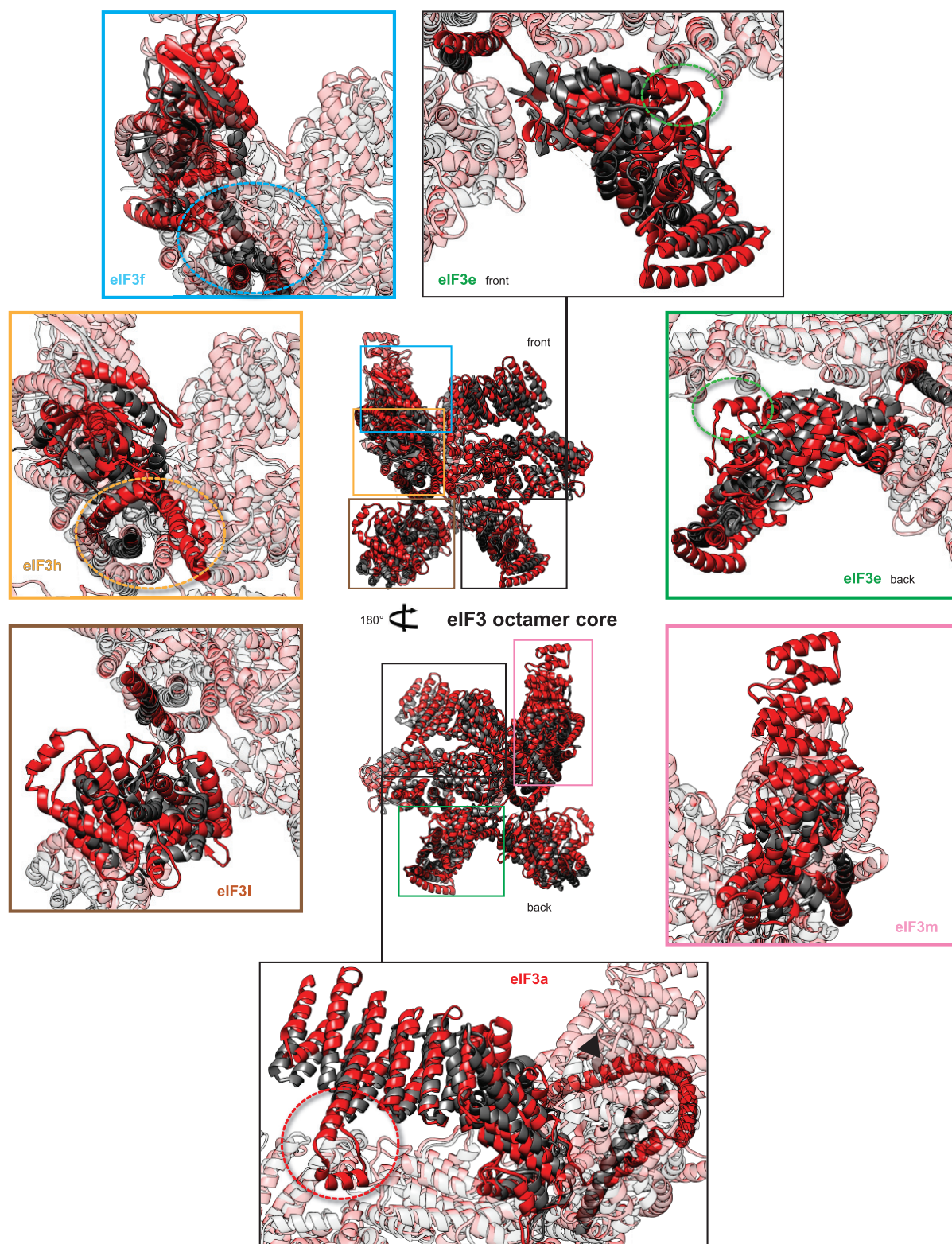
Extended Data Figure 5 | Conservation of quaternary interactions between eIF3 subunits a, m, f and h. **a**, In centre, eIF3 a, m, f and h subunits. **b–f**, Close-up views of contact regions between subunits a, m, f and h of eIF3. Red spheres represent residues at the interfaces that are conserved in six representative eukaryotic organisms; *H. sapiens*, *C. elegans*, *A. thaliana*, *D. melanogaster*, *X. tropicalis* and the unicellular *N. crassa*. Orange spheres

represent residues at the interfaces that are conserved only in the five multicellular eukaryotic organisms. The remaining residues that are suggested by the model and the density map to be involved in quaternary interactions are represented as ribbons in salmon colour. Many of these other residues are conserved in three or four of the compared organisms, but almost all of them have conserved chemical properties.



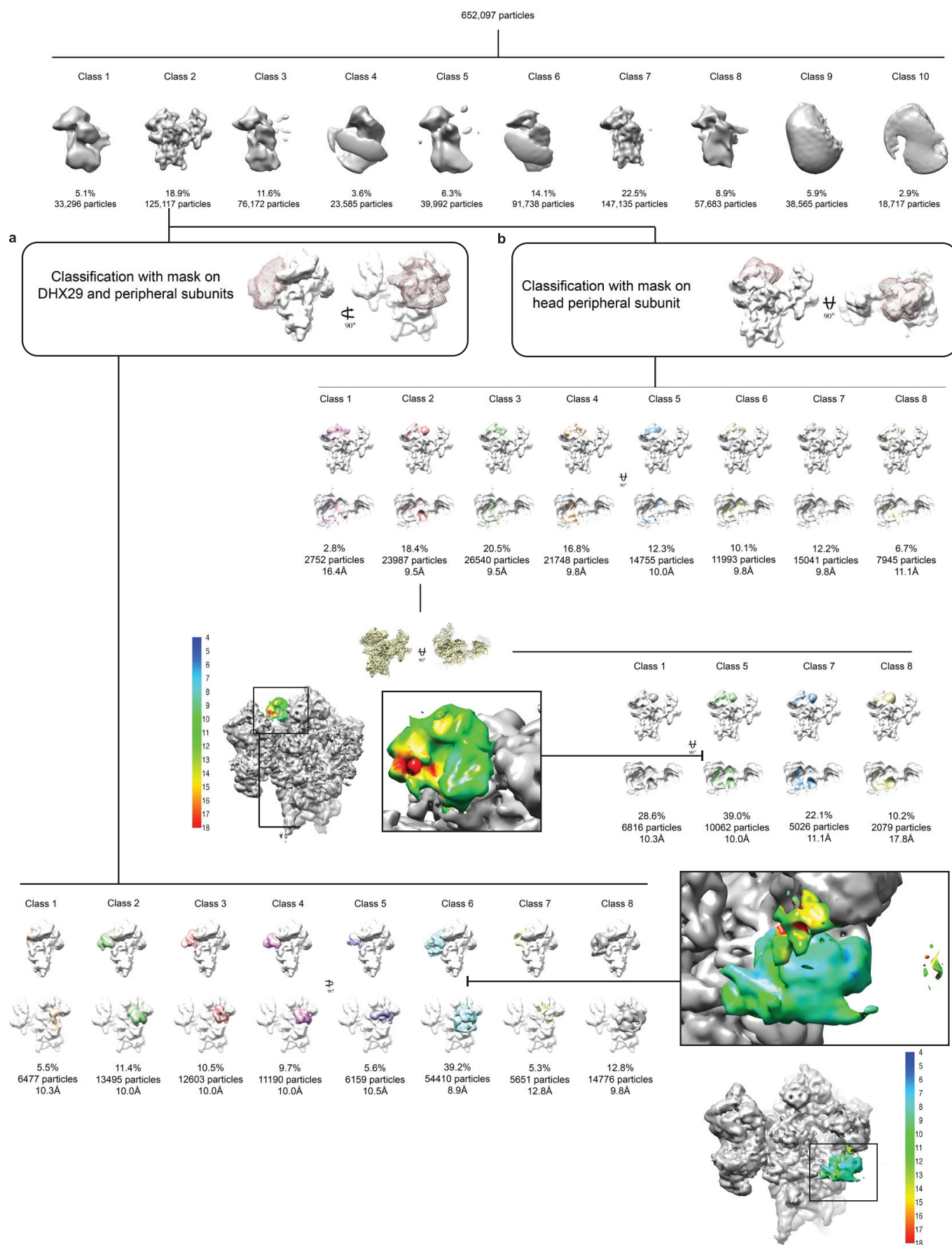
Extended Data Figure 6 | Conservation of quaternary interactions between eIF3 subunits e, k and l, and between eIF3 subunits in the region of the helical bundle. **a**, Ribbon representation of eIF3 e, k and l subunits. **b–f**, Close-up views of contact regions between subunits e, k and l of eIF3. **g**, Ribbon representation of eIF3 octamer core. **h**, Close-up views of contact regions between subunits a, c, e, m, f, h, k and l of eIF3 in the helical bundle region, seen from the direction of the axis of the latter. **i**, Region displayed in **b** rotated by 90°. Red spheres represent residues at the interfaces that are conserved in six representative eukaryotic organisms; *H. sapiens*, *C. elegans*, *A. thaliana*, *D. melanogaster*, *X. tropicalis* and the unicellular *N. crassa*. Orange spheres

represent residues at the interfaces that are conserved only in the five multicellular eukaryotic organisms. The remaining residues that are suggested by the model and the density map to be involved in quaternary interactions are represented as ribbons in salmon colour. Many of these other residues are conserved in three or four of the compared organisms, but almost all of them have conserved chemical properties. **j**, Same view as in **i** displaying all the hydrophobic residues of the helical bundle region in silver ribbons. The abundance of hydrophobic residues in the helical bundle at the interfaces of different helices suggests the stabilization of the bundle through hydrophobic interactions, hence the low identity conservation of residues at the interfaces.



Extended Data Figure 7 | Comparison of the mammalian eIF3 core model built from the 6 Å cryo-EM reconstruction with a model based on low-resolution cryo-EM studies. In the centre, our polyalanine-level model of the mammalian eIF3 octamer core (represented in red ribbons) fitted on the atomic model proposed previously¹⁸ (represented in dark grey ribbons), shown in two different orientations. The surrounding panels are close-up views of different constitutive subunits, highlighting notable structural differences between the previously proposed model¹⁸ and the model proposed in this study.

Gold- and cyan-dashed ovals highlight the misassignment of several helices of the helical bundle belonging to the C termini of subunits h and f, respectively, in the previously proposed model. Red-, green-dashed circles and black arrowhead highlight the absence in the previously proposed model of important structural features involved in quaternary interactions. In each panel, the remaining subunits of the eIF3 core octamer are faded out as transparent ribbons.



Extended Data Figure 8 | Sorting of particle images and focused classification of eIF3 peripheral subunits near the mRNA channel entrance and exit.

a. Overview of the process of sorting particle images. The population of each class is indicated by the number of particles and the percentage of the total number of particles at the beginning of each of the classification rounds. After the first round of classification, class 2 stands out as the class displaying eIF3. Focused classification of eIF3 peripheral subunits, near the mRNA channel entrance, was performed by applying a smooth-edge mask corresponding to the shape of the concerned subunits of eIF3 and to a region of the 43S complex encompassing DHX29 and h16 of the 40S subunit. The mask is displayed as pink mesh. The resulting classes from the focused sorting of class 2 of the first classification round are displayed in two different orientations, front and solvent side, respectively. Class 6 from the second classification round presents the most solid and complete density of the peripheral subunits of eIF3 at this region of the complex, and it was therefore refined yielding a reconstruction presenting an average resolution of 7.1 Å. Cryo-EM reconstruction of eIF3b and eIF3i along with DHX29, coloured by

local resolution. **b.** After the first round of classification, on class 2, focused classification of an eIF3 peripheral subunit, identified as eIF3d, near the mRNA channel exit behind ribosomal protein RACK1, was performed by applying a smooth-edge mask corresponding to the shape of the concerned subunit of eIF3 and to a region of the head of the 40S subunit encompassing RACK1. The mask is displayed as pink mesh. The resulting classes from the focused sorting of class 2 of the first classification round are displayed in two different orientations, intersubunit face and top, respectively. Class 2 from the second round of classification presents the most solid and complete density of the peripheral subunit of eIF3 at this region of the complex, but due to some apparent heterogeneity in eIF3d, a third classification round was required, yielding four classes displaying a solid eIF3d subunit in slightly different conformations (other classes obtained in this third round of classification were completely empty and therefore not shown). The major class (39% of the particles in this round) yielded a reconstruction presenting an average resolution of 7.7 Å. Cryo-EM reconstruction of eIF3d, coloured by local resolution.

Extended Data Figure 9 | Shape and ribosomal binding site of eIF3d, and eIF3b sequence. **a**, Segmented cryo-EM reconstruction of the peripheral eIF3d subunit localized on the head of the 40S subunit, behind ribosomal protein RACK1, displayed at a high density threshold to show its most solid features, in four different orientations. **b**, eIF3d subunit in the context of the 43S preinitiation complex, seen from the back, showing the ribosomal proteins that contact it directly. This figure is complementary to Fig. 4c as it displays the same complex in a different orientation. This panel shows contacts between eIF3d subunit and ribosomal proteins eS28, uS7 and uS9. Contacts with RACK1

cannot be seen from this orientation. **c**, eIF3b sequence. The amino acid sequences of human eIF3b (GenBank NP_003742.2) and rabbit eIF3b (UniProt G1SZ03_RABIT) aligned using T_COFFEE (<http://www.ebi.ac.uk/Tools/msa/tcoffee/>) and annotated to show identity with tryptic and chymotryptic peptides derived from purified rabbit eIF3 and identified by nano-LC-MS/MS analysis. The complete sequence of rabbit eIF3b has not been determined, but clearly extends beyond the N terminus of G1SZ03_RABIT, and we therefore used the numbering of residues in human eIF3b when referring in the text to elements of rabbit eIF3b.

Extended Data Table 1 | Peptide analysis of eIF3 a and c subunits after digestion and details of interactions between eIF3 and the 40S small ribosomal subunit**a**

eIF3A								
Band	Digestion with trypsin (RABBIT)				Digestion with chymotrypsin (RABBIT)			
	1st peptide	Last peptide	#Spc / #Peps	SC%	1st peptide	Last peptide	#Spc / #Peps	SC%
A	2 - 14 M.PAYFQRPENALKRA	1,348 - 1,362 R.TKNETDEDGWTTVRR-	176 / 56	28.5%	2 - 14 M.PAYFQRPENALKRA	1,348 - 1,362 R.TKNETDEDGWTTVRR-	212 / 71	31.3%
B	2 - 14 M.PAYFQRPENALKRA	1,092 - 1,100 R.NADDDRIPRR	152 / 61	30.4%	2 - 14 M.PAYFQRPENALKRA	1,101 - 1,110 R.GADDDRGFWR	149 / 59	26.0%
C	2 - 14 M.PAYFQRPENALKRA	1,092 - 1,100 R.NADDDRIPRR	201 / 02	20.0%	2 - 14 M.PAYFQRPENALKRA	1,101 - 1,110 R.GADDDRGFWR	224 / 105	27.2%
D	2 - 14 M.PAYFQRPENALKRA	1,023 - 1,030 L.DDDRGSWRT	858 / 179	49.6%	2 - 14 M.PAYFQRPENALKRA	1,050 - 1,061 RGGADDERPSWRS	1792 / 322	57.2%
E	2 - 14 M.PAYFQRPENALKRA	1,021 - 1,030 R.GLDDRGSWRT	654 / 133	44.6%	2 - 14 M.PAYFQRPENALKRA	1,021 - 1,030 R.GLDDRGSWRT	1230 / 242	50.1%
F	2 - 14 M.PAYFQRPENALKRA	980 - 990 R.RGLEDERPSWRS	168 / 53	29.7%	2 - 14 M.PAYFQRPENALKRA	1,001 - 1,010 R.LIGEDRGSWRH	262 / 84	33.5%
G	15 - 24 R.ANEFLEVGGKKQ	980 - 990 R.RGLEDERPSWRS	46 / 23	15.0%	15 - 24 R.ANEFLEVGGKKQ	1,001 - 1,010 R.LIGEDRGSWRH	70 / 32	17.6%
H	2 - 14 M.PAYFQRPENALKRA	878 - 885 R.LGEDPLSR.R	20 / 13	9.5%	2 - 14 M.PAYFQRPENALKRA	878 - 885 R.LGEDPLSR.R	23 / 14	8.4%
eIF3C								
Band	Digestion with trypsin (RABBIT)				Digestion with chymotrypsin (RABBIT)			
	1st peptide	Last peptide	#Spc / #Peps	SC%	1st peptide	Last peptide	#Spc / #Peps	SC%
A	95 - 104 K.SIVDKEGVPR.F	902 - 913 R.GGYRQQSQQTAY-	18 / 10	11.1%	95 - 104 K.SIVDKEGVPR.F	902 - 913 R.GGYRQQSQQTAY-	14 / 7	8.1%
B	95 - 104 K.SIVDKEGVPR.F	902 - 913 R.GGYRQQSQQTAY-	30 / 14	16.4%	95 - 104 K.SIVDKEGVPR.F	902 - 913 R.GGYRQQSQQTAY-	20 / 10	11.3%
C	95 - 104 K.SIVDKEGVPR.F	902 - 913 R.GGYRQQSQQTAY-	10 / 0	0.0%	95 - 104 K.SIVDKEGVPR.F	902 - 913 R.GGYRQQSQQTAY-	13 / 7	6.4%
D	95 - 104 K.SIVDKEGVPR.F	902 - 913 R.GGYRQQSQQTAY-	39 / 20	21.9%	34 - 47 K.QPLLLSEDEETKRLV	902 - 913 R.GGYRQQSQQTAY-	39 / 21	24.5%
E	95 - 104 K.SIVDKEGVPR.F	902 - 913 R.GGYRQQSQQTAY-	177 / 54	43.5%	23 - 31 L.VTKPVGGNYG	902 - 913 R.GGYRQQSQQTAY-	205 / 70	45.6%
F	2 - 33 M.SRFFTTGDSSESSLSGEEIVTK PVGGNYGKQ	902 - 913 R.GGYRQQSQQTAY-	861 / 108	56.2%	2 - 33 M.SRFFTTGDSSESSLSGEEIVTK PVGGNYGKQ	902 - 913 R.GGYRQQSQQTAY-	1219 / 204	63.5%
G	57 - 65 R.FEELTNLIR	902 - 913 R.GGYRQQSQQTAY-	65 / 23	22.8%	23 - 31 L.VTKPVGGNYG	902 - 913 R.GGYRQQSQQTAY-	79 / 34	30.2%
H	95 - 104 K.SIVDKEGVPR.F	902 - 913 R.GGYRQQSQQTAY-	369 / 52	38.8%	95 - 104 K.SIVDKEGVPR.F	902 - 913 R.GGYRQQSQQTAY-	383 / 94	42.6%

b

	eIF3a	eIF3c	eIF3b
eS1	P8,E9,L12,K13,A15,N16,L19,E20 F39,N40,I41,L74,D75,N76,D77,E 78,I189,P190 K23,K24,Q25 D191,S192,K195,D196		
eS27		L333,N334,I336,L337,I379,A380 Y383,D384,T446,L447,E449,R450 V57,G58,C59,S60,T61 G341,K342,K343 T52,V53,V54,C64,Q65,P66 R560,T561,D562,R563,D602,P603, P604,K664,V665,R668,R669 P38,G39,C40,Y41,E75,G76,C77	
eS26	P8,E9 F59,D60,A61		
uS15		N433,V434,435 K42,K43	
eS4			K390,D391,I392,P393 G118,L119,A120,P164, Y165,G166
18S	H65,L66,A67,K68 U1114,U1115 (ES7 ^s)	L389,A390,T391 G1121,U1120,A1119 (ES7 ^s) K342,K343,G344 C930,G929 ,G928 (h22)	

a. Peptide analysis of eIF3 a and c subunits, for each band of the gel the first and last identified peptides from eIF3 a and c subunits are indicated, after both types of digestions, in trypsin and chymotrypsin. The nano-LC-MS/MS analysis reveals the different forms of eIF3 a and c subunits in our *in vitro* reconstituted 43S preinitiation complex. #Spc = number of specters, #peps = number of peptides.

b. Details of interactions between eIF3 and the 40S small ribosomal subunit. The name and sequential number of interacting residues on each side, eIF3 and the 40S, is shown. Residue names are coloured variably to distinguish them according to their origin.

The formation of submillimetre-bright galaxies from gas infall over a billion years

Desika Narayanan¹, Matthew Turk², Robert Feldmann³, Thomas Robitaille⁴, Philip Hopkins⁵, Robert Thompson^{6,7}, Christopher Hayward^{5,8}, David Ball^{7,9}, Claude-André Faucher-Giguère¹⁰ & Dušan Kereš¹¹

Submillimetre-bright galaxies at high redshift are the most luminous, heavily star-forming galaxies in the Universe¹ and are characterized by prodigious emission in the far-infrared, with a flux of at least five millijanskys at a wavelength of 850 micrometres. They reside in haloes with masses about 10^{13} times that of the Sun², have low gas fractions compared to main-sequence disks at a comparable redshift³, trace complex environments^{4,5} and are not easily observable at optical wavelengths⁶. Their physical origin remains unclear. Simulations have been able to form galaxies with the requisite luminosities, but have otherwise been unable to simultaneously match the stellar masses, star formation rates, gas fractions and environments^{7–10}. Here we report a cosmological hydrodynamic galaxy formation simulation that is able to form a submillimetre galaxy that simultaneously satisfies the broad range of observed physical constraints. We find that groups of galaxies residing in massive dark matter haloes have increasing rates of star formation that peak at collective rates of about 500–1,000 solar masses per year at redshifts of two to three, by which time the interstellar medium is sufficiently enriched with metals that the region may be observed as a submillimetre-selected system. The intense star formation rates are fuelled in part by the infall of a reservoir gas supply enabled by stellar feedback at earlier times, not through major mergers. With a lifetime of nearly a billion years, our simulations show that the submillimetre-bright phase of high-redshift galaxies is prolonged and associated with significant mass buildup in early-Universe proto-clusters, and that many submillimetre-bright galaxies are composed of numerous unresolved components (for which there is some observational evidence¹¹).

We conducted our cosmological hydrodynamic galaxy formation simulations using the new hydrodynamic code GIZMO¹², which includes a model for the impact of stellar radiative and thermal pressure on the multiphase interstellar medium (ISM). This feedback both regulates the star formation rate (SFR), and shapes the structure in the ISM. Informed by clustering measurements of observed submillimetre galaxies (SMGs)², we focus on a massive halo (with a dark matter mass of $M_{\text{DM}} \approx 10^{13} M_{\odot}$ at $z = 2$, where M_{\odot} is the solar mass and z is the redshift) with baryonic particle mass $M_{\text{bary}} \approx 10^5 M_{\odot}$ as the host of our ‘main galaxy’, and run the simulation to $z = 2$. The only condition of the tracked galaxy that is pre-selected to match the physical properties of observed SMGs is the chosen halo mass. We combine this with a new dust radiation transport package, POWDERDAY, that simulates the traverse of stellar photons through the dusty ISM of the galaxy, allowing us to robustly translate our hydrodynamic simulation into observable measures. We simulate the radiative transfer from a 200 kpc region around the main galaxy. This simulation represents the first cosmological model of a galaxy this massive to be explicitly coupled with dust radiative transfer calculations. The two codes and the simulation set-up are fully described in Methods.

We define two distinct regions in the simulations. The ‘submillimetre emission region’ is the 200 kpc region surrounding the central galaxy in the halo of interest. This is the region where all of the modelled 850 μm emission comes from, and is what relates most directly to observations. The ‘submillimetre galaxy’ refers to the central galaxy in the halo. Physical quantities from the submillimetre galaxy are most applicable to high-resolution observations, as well as to placing these models in the context of other theoretical galaxy formation models. As we will show, the submillimetre emission from the region is generally dominated by the central submillimetre galaxy, though the contribution from lower mass galaxies is often non-negligible.

We track the submillimetre properties of the galaxies within the region from $z \approx 6$. The SFRs of galaxies in the region rise from this redshift towards later times $z \approx 2$, owing to accretion of gas from the intergalactic medium (Fig. 1). As stars form, stellar feedback-driven galactic winds generate outflows and fountains, allowing recycled gas to be available for star formation at later times (Extended Data Fig. 1). This phenomenon shapes a star formation history that is still rising at $z \approx 2$, in contrast to galaxy formation models with more traditional implementations of subresolution feedback, which peak at $z \approx 3$ –6 for galaxies of this mass^{13–15}. Mergers and global instabilities drive short-term variability in the global SFR, while outflows and infall driven by the feedback model can affect features in the star formation history in a somewhat cyclical ‘saw-tooth’ pattern.

At its earliest stages ($z \approx 4$ –6), the integrated SFR from the galaxies in the region varies in the approximate range $(100\text{--}300) M_{\odot} \text{ yr}^{-1}$, with a significant stellar mass, $(0.5\text{--}1) \times 10^{11} M_{\odot}$, in place, comparable to some high-redshift detections¹⁶. Feedback from massive stars enriches the ISM with metals, and the dust content simultaneously rises. By $z \approx 3$, the combination of gas accumulation and substantial metal enrichment drives an increase in the dust mass by a factor of ~ 50 , with masses approaching $\sim 1 \times 10^9 M_{\odot}$. Radiation from the delayed peak in the SFR interacting with this substantive dust reservoir drives the observed 850 μm flux density to detectable values of $> 5 \text{ mJy}$. The galaxies associated with the main halo enter a long-lived submillimetre-luminous phase, with a lifetime of $\sim 0.75 \text{ Gyr}$. While our main model is only run to $z = 2$ owing to computational restrictions for models of this resolution, tests with lower-resolution models reveal that at later times ($z < 1.5$), a declining SFR due to inefficient accretion as well as exhausted gas supply drives a drop in the submillimetre flux density (for more details, see Methods). The star formation history of galaxies residing in haloes with $M_{\text{DM}} \approx 10^{13} M_{\odot}$ (at $z = 2$), as controlled by the underlying stellar feedback, provides a physical explanation for the peak in the observed SMG redshift distribution at $z = 2$ –3 (ref. 17).

During the submillimetre-luminous phase, the emitting region is almost always occupied by multiple detectable galaxies. In Fig. 2, we

¹Haverford College, 370 West Lancaster Avenue, Haverford, Pennsylvania 19041, USA. ²National Center for Supercomputing Applications, University of Illinois, 1205 West Clark Street, Urbana-Champaign, Illinois 61820, USA. ³Department of Astronomy and Theoretical Astrophysics Center, University of California, Berkeley, California 94720, USA. ⁴Max Planck Institute for Astronomy, Königstuhl 17, D-69117 Heidelberg, Germany. ⁵TAPIR, California Institute of Technology, MC 350-17, Pasadena, California 91125, USA. ⁶University of the Western Cape, Bellville, Cape Town 7535, South Africa. ⁷Steward Observatory, University of Arizona, 933 North Cherry Avenue, Tucson, Arizona 85721, USA. ⁸Harvard-Smithsonian Center for Astrophysics, 60 Garden Street, Cambridge, Massachusetts 02138, USA. ⁹Whitman College, 345 Boyer Avenue, Walla Walla, Washington 99362, USA. ¹⁰CIERA, Northwestern University, 2145 Sheridan Road, Evanston, Illinois 60208, USA. ¹¹CASS, University of California, San Diego, 9500 Gilman Drive, La Jolla, California 92093, USA.

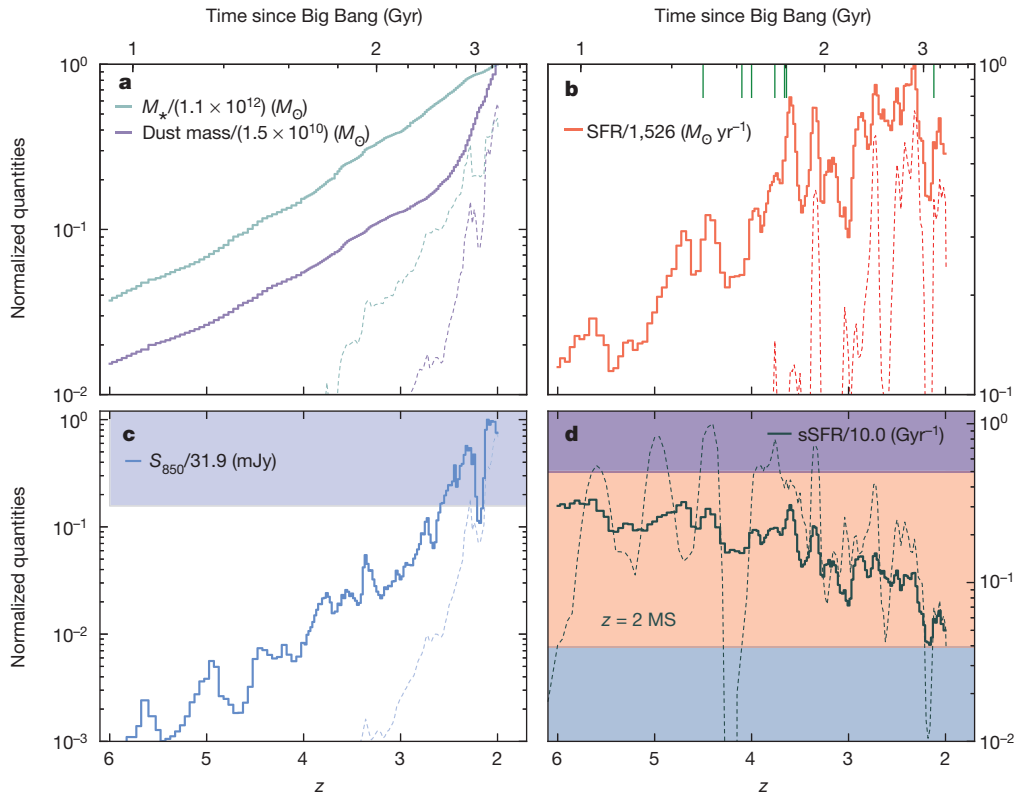


Figure 1 | Evolution of physical and observable properties of the submillimetre emission region and the central galaxy. In each panel, the properties of the 200 kpc submillimetre emission region are shown with thick solid lines, while those of the central galaxy are given by thin dashed lines. **a**, Stellar and dust mass; **b**, SFR; **c**, predicted observed 850 μm flux density; **d**, specific SFR, $s\text{SFR}$ (M_*/SFR). The SFR is averaged on 50 Myr timescales, and

includes a correction factor of 0.7 for mass loss. Locations of major galaxy mergers ($>1:3$) are noted by green vertical ticks on the top axis of **b**. The light purple shaded region in **c** shows when the galaxy would be detectable as an SMG with SCUBA ($S_{850} > 5$ mJy). The pink and purple shaded regions in **d** show the rough ranges for the $z = 2$ main-sequence (MS) and starburst regime; the grey region in **d** denotes below the main sequence.

present gas surface density projections of six arbitrarily chosen snapshots during the evolution of the submillimetre-luminous phase ($z = 2-3$). The panels are 250 kpc on a side; for reference, the full-width at half-maximum (FWHM) of the Submillimetre Common-user Bolometer Array (SCUBA) on the James Clerk Maxwell Telescope, the first instrument to detect SMGs, is ~ 125 kpc at $z \approx 2$. Multiple clumps of gas falling into the central galaxy are nearly always present. The observed flux density from the region is typically dominated by the central galaxy, with (on average) $\sim 30\%$ arising from emission from subhaloes (Extended Data Fig. 2). The submillimetre flux density of the central galaxy rises dramatically between $z \approx 2-3$, reaching a peak value of ~ 20 mJy. Owing to contributions from subhaloes surrounding the central galaxy, the flux from the overall 200 kpc region can exceed this, peaking at ~ 30 mJy. Similarly extreme systems have recently been detected with the Herschel Space Observatory and the South Pole Telescope^{4,18,19}.

While the central galaxy is being bombarded by subhaloes over a range of mass ratios during the submillimetre-luminous phase, major galaxy mergers akin to local prototypical analogues such as Arp 220 or NGC 6240 do not drive the onset of the long-lived submillimetre-luminous phase in the central galaxy. In Fig. 1, we highlight when the galaxy undergoes a major merger with mass ratio $\geq 1:3$. While major mergers are common at early times (and indeed drive some short-lived bursts in star formation), the bulk of the submillimetre-luminous phase at later times ($z \approx 2-3$) occurs nearly a gigayear after the last major merger. The ratio of the SFR to its integral over cosmic time (the specific SFR) of the overall emitting region is generally on the main sequence of galaxy formation at $z \approx 2$ (defined as the main locus of points on the SFR–stellar mass (M_*) relation), although the central galaxy can have values comparable both to main-sequence galaxies

between $z \approx 2-3$ and to outliers. One consequence of a model in which SMGs typically lie on the main sequence of star formation is that the gas surface densities show a broad range, $\sim (10^2-10^4) M_\odot \text{pc}^{-2}$ (Extended Data Fig. 3), as well as diverse gas spatial extents (Fig. 3). This is manifested observationally in the broad swath occupied by SMGs in the Kennicutt–Schmidt star formation relation¹. The spatial extent and surface density of the gas are to be contrasted, however, with local merger-driven ultraluminous infrared galaxies, which exhibit typical FWHM radii of $\sim 100-500$ pc (ref. 20). Idealized galaxy merger simulations with initial conditions designed to form SMGs further underscore this contrast, as they also result in compact morphologies during final coalescence, and can be inefficient producers of submillimetre radiation owing to increased dust temperatures⁸.

The central submillimetre galaxy is amongst the most massive and highly star-forming of galaxies at this epoch. The stellar masses are diverse, in the range $\sim (1-5) \times 10^{11} M_\odot$, comparable to recent measurements of this population²¹, and consistent with constraints from abundance matching techniques²². The molecular gas fractions of the central galaxy ($f_{\text{gas}} \equiv M_{\text{H}_2}/(M_{\text{H}_2} + M_*)$) decline with stellar mass, and range from $\sim 40\%$ at lower stellar masses to $\lesssim 10\%$ at the highest masses. This is in agreement with observations²³, although is dependent on the conversion from carbon monoxide (^{12}CO) luminosity to H_2 gas mass. We note that these predictions are quantitatively different from those produced by previous cosmological efforts in this field, with some predicted gas fractions exceeding $f_{\text{gas}} = 0.75$ (refs 7, 9) and median stellar masses as low as $\sim 10^{10} M_\odot$ (ref. 7). We present plots of the gas fractions and calculated spectral energy distributions (SEDs) of our model SMG in the context of observations in Extended Data Figs 4 and 5. The modelled gas distributions within the central galaxy, which

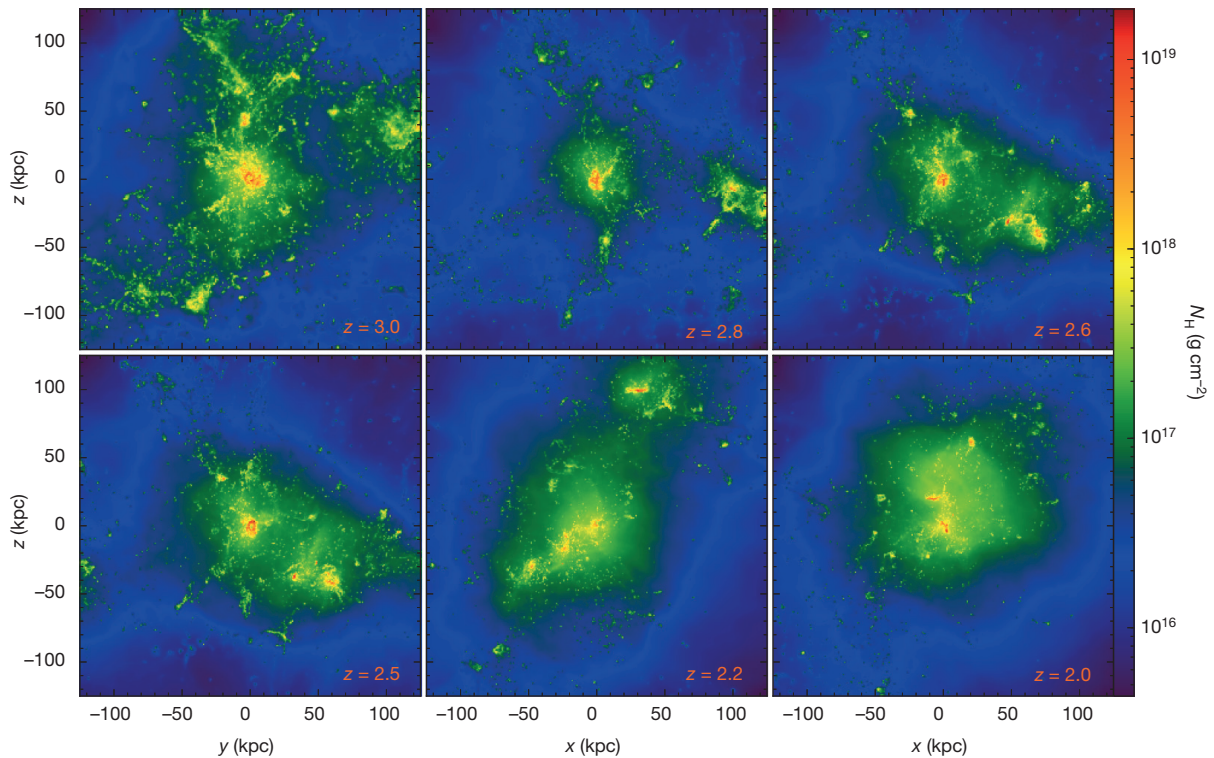


Figure 2 | Surface density projection maps of the 250 kpc region around the central submillimetre galaxy for redshifts $z \approx 2$ –3. The submillimetre emission region probed in surveys typically encompasses a central galaxy in a massive halo that is undergoing a protracted bombardment phase by numerous

subhaloes. Some of the brightest SMGs arise from numerous galaxies within the beam in a rich environment (bottom right panel). The colour coding denotes the gas column density (N_{H}), with the colour bar on the right.

range from ~ 1 kpc to 8 kpc, compare well with recent dust maps observed using the Atacama Large Millimetre Array²⁴.

The stellar masses, gas fractions and lifetimes are in agreement with some previous lower-resolution cosmological efforts¹⁰, although the predicted SFR and luminosity from this model are substantially larger. The SFR of the group of galaxies in the region peaks at $\sim 1,500 M_{\odot} \text{ yr}^{-1}$. Importantly, up to half of the total infrared luminosity

can come from older stars with ages $t_{\text{age}} > 0.1$ Gyr. Using standard conversions²⁵, the estimated SFR from the integrated infrared SED (3 – $1,100 \mu\text{m}$) can exceed $\sim 3,000 M_{\odot} \text{ yr}^{-1}$ (Extended Data Fig. 6), and hence infrared-based SFR derivations of dusty galaxies at high z may over-estimate the true SFR by a factor of ~ 2 . Indeed, the contribution of satellite galaxies to the global SFR, along with the contribution of old stars to the infrared luminosity may relieve some tensions between the inferred SFRs from submillimetre galaxies and massive galaxies modelled in cosmological hydrodynamic simulations¹⁰.

The end-product of the central submillimetre galaxy at $z \approx 2$ is a galaxy with a stellar mass of $\sim (4\text{--}5) \times 10^{11} M_{\odot}$ that is distributed over a compact region of ~ 1 – 5 kpc, and gas that is distributed similarly (Fig. 3). This is similar in extent and mass to the observed $z \approx 2$ compact quiescent galaxy population, which has a mean half-light radius of $R_e \approx 1.5$ kpc, a stellar mass of $M_* > 10^{11} M_{\odot}$ and ages $t_{\text{age}} \approx 0.5$ – 1 Gyr (ref. 26), suggesting a plausible connection between the galaxy populations. Indeed, a calculation of the stellar velocity dispersion along three orthogonal sightlines of the central galaxy during the submillimetre-luminous phase results in $\sigma_* \approx 600$ – 700 km s^{-1} , comparable to measurements of high- z compact quiescents. A large sample of simulated SMGs would allow for a robust analysis of the expected abundances of SMGs and compact quiescents at $z \approx 2$.

Our picture for the formation of SMGs suggests that they are not transient events, but rather natural long-lived phases in the evolution of massive haloes. The ~ 0.75 Gyr duty cycle combined with the comoving abundance²⁷ of dark matter haloes of this mass result in an expected abundance of our model SMGs of $\sim 1.5 \times 10^{-5} h^3 \text{ Mpc}^{-3}$, comparable to the $\sim 10^{-5} h^3 \text{ Mpc}^{-3}$ observed for SMGs²⁸. While modelling the full number counts involves convolving the typical duty cycle as a function of halo mass with halo mass functions over a range of redshifts, the approximate abundances implied by this model are encouraging.

This model suggests that galaxies that form in haloes of mass $M_{\text{DM}} \approx 10^{14} M_{\odot}$ at $z = 0$ will represent typical SMGs near the peak of their redshift distribution. Lower mass galaxy models do not achieve

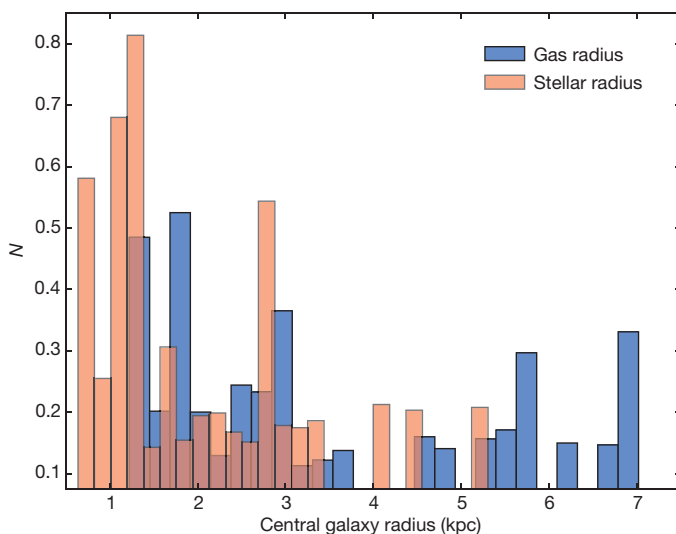


Figure 3 | Gas and stellar radius distribution for the central submillimetre galaxy. The orange histogram denotes the half-mass radius of the stars, while the blue shows the gas. The galaxy gas is more distributed in the central galaxy than the (subkiloparsec) extent expected from major mergers, although still sufficiently compact that it will remain unresolved even with approximately arcsecond resolution. The ordinate is weighted by the time the galaxy spends in the bin, and the overall normalization of the distribution is arbitrary.

the requisite SFR and metal enrichment to generate submillimetre-luminous galaxies (see Methods). More extreme SMGs that are being detected for $z = 5\text{--}6$ (refs 29, 30) may form in even more massive (and rare) haloes than those considered here.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 25 October 2014; accepted 31 July 2015.

- Casey, C. M., Narayanan, D. & Cooray, A. Dusty star-forming galaxies at high redshift. *Phys. Rep.* **541**, 45–161 (2014).
- Hickox, R. C. *et al.* The LABOCA survey of the Extended Chandra Deep Field-South: clustering of submillimetre galaxies. *Mon. Not. R. Astron. Soc.* **421**, 284–295 (2012).
- Geach, J. E. *et al.* On the evolution of the molecular gas fraction of star-forming galaxies. *Astrophys. J.* **730**, L19 (2011).
- Fu, H. *et al.* The rapid assembly of an elliptical galaxy of 400 billion solar masses at a redshift of 2.3. *Nature* **498**, 338–341 (2013).
- Daddi, E. *et al.* Two bright submillimetre galaxies in a $z = 4.05$ protocluster in GOODS-North, and accurate radio-infrared photometric redshifts. *Astrophys. J.* **694**, 1517–1538 (2009).
- Swinbank, A. M. *et al.* The rest-frame optical spectra of SCUBA galaxies. *Astrophys. J.* **617**, 64–80 (2004).
- Baugh, C. M. *et al.* Can the faint submillimetre galaxies be explained in the Λ cold dark matter model? *Mon. Not. R. Astron. Soc.* **356**, 1191–1200 (2005).
- Hayward, C. C. *et al.* Submillimetre galaxies in a hierarchical universe: number counts, redshift distribution and implications for the IMF. *Mon. Not. R. Astron. Soc.* **428**, 2529–2547 (2013).
- Shimizu, I., Yoshida, N. & Okamoto, T. Submillimetre galaxies in cosmological hydrodynamic simulations: source number counts and the spatial clustering. *Mon. Not. R. Astron. Soc.* **427**, 2866–2875 (2012).
- Davé, R. *et al.* The nature of submillimetre galaxies in cosmological hydrodynamic simulations. *Mon. Not. R. Astron. Soc.* **404**, 1355–1368 (2010).
- Hodge, J. A. *et al.* An ALMA survey of submillimetre galaxies in the extended Chandra Deep Field South: source catalog and multiplicity. *Astrophys. J.* **768**, 91 (2013).
- Hopkins, P. F. GIZMO: a new class of accurate, mesh-free hydrodynamic simulation methods. *Mon. Not. R. Astron. Soc.* **450**, 53–110 (2015).
- Davé, R., Finlator, K. & Oppenheimer, B. D. An analytic model for the evolution of the stellar, gas and metal content of galaxies. *Mon. Not. R. Astron. Soc.* **421**, 98–107 (2012).
- Hopkins, P. F. *et al.* Galaxies on FIRE (Feedback In Realistic Environments): stellar feedback explains cosmologically inefficient star formation. *Mon. Not. R. Astron. Soc.* **445**, 581–603 (2014).
- Feldmann, R. & Mayer, L. The Argo Simulation—I. Quenching of massive galaxies at high redshift as a result of cosmological starvation. *Mon. Not. R. Astron. Soc.* **446**, 1939–1956 (2015).
- Finkelstein, S. L. *et al.* A galaxy rapidly forming stars 700 million years after the Big Bang at redshift 7.51. *Nature* **502**, 524–527 (2013).
- Weiβ, A. *et al.* ALMA redshifts of millimeter-selected galaxies from the SPT survey: the redshift distribution of dusty star-forming galaxies. *Astrophys. J.* **767**, 88 (2013).
- Ivison, R. J. *et al.* Herschel-ATLAS: a binary HyLIRG pinpointing a cluster of starbursting protoellipticals. *Astrophys. J.* **772**, 137 (2013).
- Hezaveh, Y. D. *et al.* ALMA observations of SPT-discovered, strongly lensed, dusty, star-forming galaxies. *Astrophys. J.* **767**, 132 (2013).
- Downes, D. & Solomon, P. M. Rotating nuclear rings and extreme starbursts in ultraluminous galaxies. *Astrophys. J.* **507**, 615–654 (1998).
- Michałowski, M. J. *et al.* The stellar masses and specific star-formation rates of submillimetre galaxies. *Astron. Astrophys.* **541**, A85 (2012).
- Behroozi, P. S., Wechsler, R. H. & Conroy, C. The average star formation histories of galaxies in dark matter halos from $z=0\text{--}8$. *Astrophys. J.* **770**, 57 (2013).
- Tacconi, L. J. *et al.* Pibss: Molecular gas content and scaling relations in $z \sim 1\text{--}3$ massive, main-sequence star-forming galaxies. *Astrophys. J.* **768**, 74 (2013).
- Simpson, J. M. *et al.* The SCUBA-2 cosmology legacy survey: ALMA resolves the rest-frame far-infrared emission of sub-millimeter galaxies. *Astrophys. J.* **799**, 81 (2015).
- Kennicutt, R. C. & Evans, N. J. Star formation in the Milky Way and nearby galaxies. *Annu. Rev. Astron. Astrophys.* **50**, 531–608 (2012).
- van Dokkum, P. G. *et al.* Confirmation of the remarkable compactness of massive quiescent galaxies at $z \sim 2.3$: early-type galaxies did not form in a simple monolithic collapse. *Astrophys. J.* **677**, L5–L8 (2008).
- Murray, S. G., Power, C. & Robotham, A. S. G. HMFcalc: an online tool for calculating dark matter halo mass functions. *Astron. Comput.* **3–4**, 23–34 (2013).
- Chapman, S. C., Blain, A. W., Smail, I. & Ivison, R. J. A redshift survey of the submillimetre galaxy population. *Astrophys. J.* **622**, 772–796 (2005).
- Riechers, D. A. *et al.* A dust-obscured massive maximum-starburst galaxy at a redshift of 6.34. *Nature* **496**, 329–333 (2013).
- Vieira, J. D. *et al.* Dusty starburst galaxies in the early Universe as revealed by gravitational lensing. *Nature* **495**, 344–347 (2013).

Acknowledgements We thank M. J. Michałowski for providing observational data. Partial support for D.N. was provided by NSF AST-1009452, AST-1442650, NASA HST AR-13906.001 and a Cottrell College Science Award. P.H., C.H., M.T. and R.T. were funded by the Gordon and Betty Moore Foundation (GBMF4561 and grant no. 776). P.H. acknowledges the Alfred P. Sloan Foundation for support. C.-A.F.-G. was supported by NASA awards PF3-140106, NNX15AB22G and NSF AST-1412836. D.K. was supported by NSF AST-1412153. R.F. was supported by NASA HF-51304.01-A, and is a Hubble fellow. The simulations here were run on Stampede at TACC through NSF XSEDE allocations TG-AST120025, TG-AST130039 and TG-AST140023, NASA Pleiades, and the Haverford College cluster.

Author Contributions D.N. wrote the text, and led the radiative transfer simulations and analysis. D.N., M.T., T.R. and R.T. wrote the POWDERDAY software. R.T., C.H. and D.B. contributed to simulation analysis, and R.F., P.H., C.-A.F.-G. and D.K. performed the cosmological simulations.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to D.N. (dnarayan@haverford.edu).

METHODS

Cosmological hydrodynamic zoom simulations. We utilize a newly developed version of TreeSPH that employs a pressure–entropy formulation of smoothed particle hydrodynamics (SPH)³¹ that obviates many of the potential discrepancies noted between grid-based codes, traditional SPH codes, and moving-mesh algorithms^{32–34}. In particular, we employ the hydrodynamic code GIZMO¹² in P-SPH mode which conserves momentum, energy, angular momentum and entropy, and includes newly developed algorithms to treat the artificial viscosity, entropy diffusion and time-stepping^{31,35}. The gravity solver is a modified version of the GADGET-3 solver³⁶, and an updated softening kernel to better represent the potential of the SPH smoothing kernel is included³⁷.

The simulations are fully cosmological zoom-in calculations of the evolution of individual galaxies. A 144 Mpc³ cosmological volume was simulated at low resolution down to redshift $z = 0$ with dark matter only. The halo of interest was identified, and re-simulated at much higher resolution with baryons included. The initial conditions were generated with the MUSIC code³⁸. We simulate four zoom galaxies—one is our main galaxy, and the other three are at varying resolutions and masses for the purposes of testing. The main galaxy of interest to this study resides in a dark matter halo mass of $M_{\text{DM}} = 3 \times 10^{13} M_{\odot}$ at $z = 2$. The initial baryonic particle masses in the high-resolution region were $2.7 \times 10^5 M_{\odot}$, and the minimum baryonic/stars/dark matter force softening lengths were 9/21/142 proper pc at $z = 2$. The physical properties of all of the modelled galaxies are presented in Extended Data Table 1.

The baryonic physics implemented into GIZMO are developed on the basis of extensive tests studying idealized simulations of both isolated disks and galaxy mergers^{39–44}. The gas cools using an updated cooling curve to standard⁴⁵ implementations in SPH codes which includes both atomic and molecular line emission⁴⁶. The modelled ISM is multiphase. The neutral ISM is broken into an atomic and molecular component following algorithms that scale the molecular fraction with column density and gas phase metallicity^{47,48}. Star formation occurs in molecular gas above a threshold density (here, this is set to $n_{\text{thresh}} = 10 \text{ cm}^{-3}$). Star formation is further restricted to gas that is locally self-gravitating, where:

$$\alpha \approx \beta' \frac{|\nabla \cdot \mathbf{v}|^2 + |\nabla \times \mathbf{v}|^2}{G\rho} < 1 \quad (1)$$

where α is the usual virial parameter, $\beta' \approx 1/2$, \mathbf{v} the gas velocity, G is the gravitational constant and ρ is the density of the gas. This follows from studies⁴⁹ that show that the predicted spatial distribution of star formation in galaxies is more realistic when using a gas self-gravitating criterion compared to a variety of other algorithms (including a fixed density threshold, a pure molecular-gas law, a temperature threshold, a Jeans criterion, a cooling-time criterion and a converging flow criterion). The SFR follows a volumetric relation:

$$\dot{\rho}_* = \rho_{\text{mol}}/t_{\text{ff}} \quad (2)$$

where $\dot{\rho}_*$ is the star formation rate density, ρ_{mol} the gas density, and t_{ff} the local gas free fall time. In other words, stars are allowed to form with 100% efficiency per free fall time. The star formation is subsequently self-regulated by stellar feedback, resulting in a time-averaged efficiency on galaxy scales of ϵ_{ff} of ~ 0.005 – 0.1 (ref. 39).

Once stars have formed, they can impact the ISM via various feedback mechanisms. Assuming a Kroupa⁵⁰ stellar initial mass function, and using STARBURST99⁵¹ for luminosity, mass-return and supernova rate calculations as a function of stellar age and metallicity, we include the following forms of stellar feedback.

Radiation momentum deposition. At each timestep, the gas near young stars is affected by a momentum flux given by:

$$\dot{p}_{\text{rad}} \approx (1 - \exp(-\tau_{\text{UV/optical}}))(1 + \tau_{\text{IR}})L_{\text{incident}}/c \quad (3)$$

where L_{incident} is the incident luminosity, $\tau_{\text{UV/optical}}$ is the optical depth to UV/optical photons, $\tau_{\text{IR}} = \Sigma_{\text{gas}}\kappa_{\text{IR}}$, Σ_{gas} is the column of gas and $\kappa_{\text{IR}} = 5(Z/Z_{\odot})g^{-1} \text{ cm}^2$.

Supernovae and stellar winds. We utilize tabulated type-I and type-II supernovae rates^{51,52}; if a supernova occurs during a timestep, thermal energy and radial momentum are injected within a smoothing length of the star. Gas and metal return are included as well. Stellar winds are similarly included with energy, wind momentum, mass and metals deposited within a smoothing length.

Photoheating of H II regions. The production rate of ionizing radiation from stars determines the extent of H II regions (allowing for overlapping regions). These regions are heated to 10^4 K if the gas is below that threshold.

We utilize models TL37 SR and TL37 HR in Extended Data Table 1 to test the convergence properties of our simulations. One model is run with the same mass baryonic resolution as our main model (standard resolution; SR), and one a factor of ~ 8 higher resolution (high resolution; HR). In Extended Data Fig. 7, we show the modelled duty cycle above a given flux density as a function of flux density for these two models. We see that the shortest lived ($\lesssim 200 \text{ Myr}$) emission spikes

present in the standard resolution model may not be converged in the highest resolution model. Notably, emission with longer duty cycles is either converged, or underpredicted in our standard resolution model, suggesting that the relatively long-lived submillimetre-luminous phase is robust.

We show the M_* – z relation for the central galaxy in Extended Data Fig. 8 as compared to observational constraints²². The central galaxy has a stellar mass a factor of ~ 2 greater than the observed median stellar mass for comparable mass haloes at this epoch. The model galaxy may represent an outlier in the M_* – z relation at these redshifts. Indeed, the thickness of the observational constraints shows the uncertainty, not range, in possible values. Alternatively, it is possible that the inclusion of feedback from an active galactic nucleus (AGN) could impact the stellar mass buildup in the galaxy, although the level to which black hole growth can impact star formation near the submillimetre-luminous phase is unclear. Some models have shown that AGN can grow efficiently in the absence of major mergers^{53,54}, while other models and observations suggest that mergers may be necessary to grow massive holes^{55–57}. The last major merger before the submillimetre-luminous phase is $\sim 1 \text{ Gyr}$ before. Tests with our low resolution model (m13m14) show that without AGN feedback, residual star formation drives a factor of ~ 2 increase in stellar mass at late times ($z \approx 0$ – 1). Finally, we note that a higher mass resolution model could potentially also result in decreased final stellar masses. In our convergence tests, the final M_* (at $z = 2$) of the HR run is $\sim 60\%$ that of the SR run.

Dust radiative transfer calculations. To calculate the inferred observational properties of our simulated galaxies, we developed a dust radiative transfer package, POWDERDAY. In short, POWDERDAY takes hydrodynamic simulations of galaxies in evolution, projects the gas properties onto an adaptive mesh and calculates the radiative transfer from the stellar sources through the dusty ISM until an equilibrium dust temperature is achieved.

In detail, we identify galaxies using SKID to locate bound groups of baryonic particles^{58,59}, and track their progenitors back in time^{60,61}. Galaxies and haloes are required to contain at least 64 particles each in order to be identified. We cut out a 200 kpc (side length) region around the galaxy of interest, and subdivide the domain into an adaptive grid with an octree memory structure. Formally, we begin with one cell encompassing the entire $8 \times 10^6 \text{ kpc}^3$ radiative transfer region. The cells then recursively subdivide into octs until there are a threshold maximum number of gas particles in the cell (we employ $n_{\text{subdivide,thresh}} = 64$, although experiments with $n_{\text{subdivide,thresh}} = 32$ show converged results). The physical properties of the gas particles are projected onto the octree using a spline smoothing kernel⁶².

The spectral energy distribution of stars are calculated on the fly with the Flexible Stellar Population Synthesis code, FSPS^{63,64}, through PYTHON-FSPS, a set of PYTHON hooks for FSPS (<https://github.com/dfm/python-fsps>). The SEDs are calculated as simple stellar populations with ages and metallicities determined by the hydrodynamic simulation, and assuming a Kroupa IMF.

The radiative transfer happens in a Monte Carlo fashion using the three-dimensional dust radiative transfer solver, HYPERION⁶⁵. The code uses an iterative methodology to determine the radiative equilibrium temperature⁶⁶, and we determine convergence when the energy absorbed by 99% of the cells has changed by less than 1% between iterations. We assume a dust grain-size distribution comparable to that of the Milky Way⁶⁷, with $R \equiv A_v/E(B - V) = 3.15$, where A_v is the visual extinction and $E(B - V)$ is the difference between the B- and V-band extinctions. The dust emissivities are updated to include an approximation for polycyclic aromatic hydrocarbons (PAHs) alongside thermal emission⁶⁸. We assume a constant dust to metals ratio of 0.4, motivated by both Milky Way and extragalactic observational constraints^{69–71}.

The underlying HYPERION code has passed the standard benchmarks for codes of this type⁷², and we found that POWDERDAY compares well against other publicly available dust radiative transfer codes^{73–75} in test starburst SPH galaxy merger simulations.

Parameter choices. In Extended Data Fig. 9, we present a number of tests of our parameter choices for the radiative transfer calculations. We show the predicted 850 μm light curve from our lowest resolution model (m13m14) using fiducial parameters, as well as three parameter choice variations.

We first ask whether our chosen radiative transfer grid size affects our principal results. Our fiducial model is a 200 kpc (on a side) box cut out of the global cosmological simulation centred on the halo of interest. This size was chosen to reflect a rough average of the typical (sub)millimetre beam sizes typically used to detect SMGs. For example, assuming Planck 2013 cosmological parameters⁷⁶, the Submillimetre Common-Use Bolometer Array (SCUBA) on the James Clerk Maxwell Telescope (JCMT) has a $15''$ full-width at half-maximum (FWHM) beam at 850 μm . At $z = 2$ this corresponds to $\sim 128 \text{ kpc}$. At the same redshift, the beam of AzTEC and LABOCA at 1 mm on the JCMT corresponds to $\sim 163 \text{ kpc}$ ($19''$); the South Pole Telescope (SPT) has a beamsize of 540 kpc at 1.4 mm ($63''$); and Herschel's SPIRE instrument ranges from 154 kpc to 308 kpc (250–500 μm ; $18''$ – $36''$).

Because a few notable beam sizes (of particular relevance, the SCUBA beam) are smaller than our assumed box size of 200 kpc, we have run an additional model with box length of 100 kpc (and all other parameters exactly the same). We highlight the resultant 850 μm light curve from this model in the top right panel of Extended Data Fig. 9. When comparing to our fiducial model, it is apparent that our results are robust to the highest resolution beams that have been used for SMG surveys at single dish facilities to date.

We additionally investigate whether our inclusion of PAHs in our model makes any difference to the calculated submillimetre-wave flux density of our model galaxy. This is presented in the bottom left panel of Extended Data Fig. 9. Again, we note minimal impact on the submillimetre SED of our model.

Finally, we ensure that our results are converged with the number of photons emitted. We fiducially run 10^7 photons per grid (roughly 100 per cell). In the bottom right panel of Extended Data Fig. 9, we show the results from a run with 10^8 photons per grid, and show that the results are robust against this parameter choice.

Relation to other models. Historically, the methods used, and physical models for SMG formation in numerical simulations are quite varied. Here, we summarize these methods and results, and place our own model into this context. Broadly, there are three classes of SMG formation models: cosmological semi-analytic models (SAMs), idealized non-cosmological simulations and cosmological hydrodynamic models. The present model falls into the last category. Our model is the first self-consistent cosmological simulation with baryons and bona fide radiative transfer to form a submillimetre galaxy with physical properties comparable to those observed.

The initial forays into this field were typically with SAMs. This is because SAMs are computationally inexpensive, and allow for a large search in physical parameter space relatively easily. SAMs either utilize analytic halo merger trees, or directly simulate them, and then employ analytic prescriptions to describe the central galaxies. The Durham SAM^{7,77} couples galaxies formed in a semi-analytic model with dust radiative transfer. These simulations model galaxies that have axisymmetric geometries that consist of a disk and a bulge. Young stellar populations are assumed to still be enveloped in their birth clouds, and thus experience additional attenuation. This model suggests that roughly $\sim 22\%$ of SMGs originate from major mergers, the remainder from minor mergers, and that the stellar IMF is flat during the starburst. The typical lifetime for the submillimetre-luminous phase is ~ 100 Myr (a factor of ~ 7.5 lower than found in our work), galaxies are extremely gas rich ($f_{\text{gas}} \approx 75\%$), and stellar masses are a factor of ~ 10 lower than predicted by our model ($M_* \approx 2 \times 10^{10} M_\odot$). While the stellar masses of SMGs are debated^{42,78,79}, the gas fractions appear to be uniformly lower in observations^{6,18,23,80,81}, and a flat stellar IMF is probably ruled out by CO dynamical mass measurements⁸².

As an alternative to cosmological simulations, a number of studies have explored SMG formation in idealized simulations^{8,83–86}. These studies evolve hydrodynamic models of idealized disks and mergers over a range of merger mass ratios, and combine these with dust radiative transfer simulations⁷⁴. These models infer halo masses and stellar masses for SMGs comparable to those modelled here. This said, in the idealized galaxy models, $\sim 30\%$ – 70% of the SMGs (flux dependent) originate in merger-driven starbursts, substantially higher than what is found for our model. Some studies⁸ have noted that binary mergers that cause SMGs may break up into multiples at high-resolution owing to the contribution to the total flux of individual inspiralling disks. Because idealized simulations are non-cosmological in nature, comparing the multiplicity inferred from these to our models is difficult: the major merger multiplicity can only be two when considering galaxies at the same redshift. On the other hand, Extended Data Fig. 2 suggests that potentially larger multiplicity can be observed for physically associated clumps.

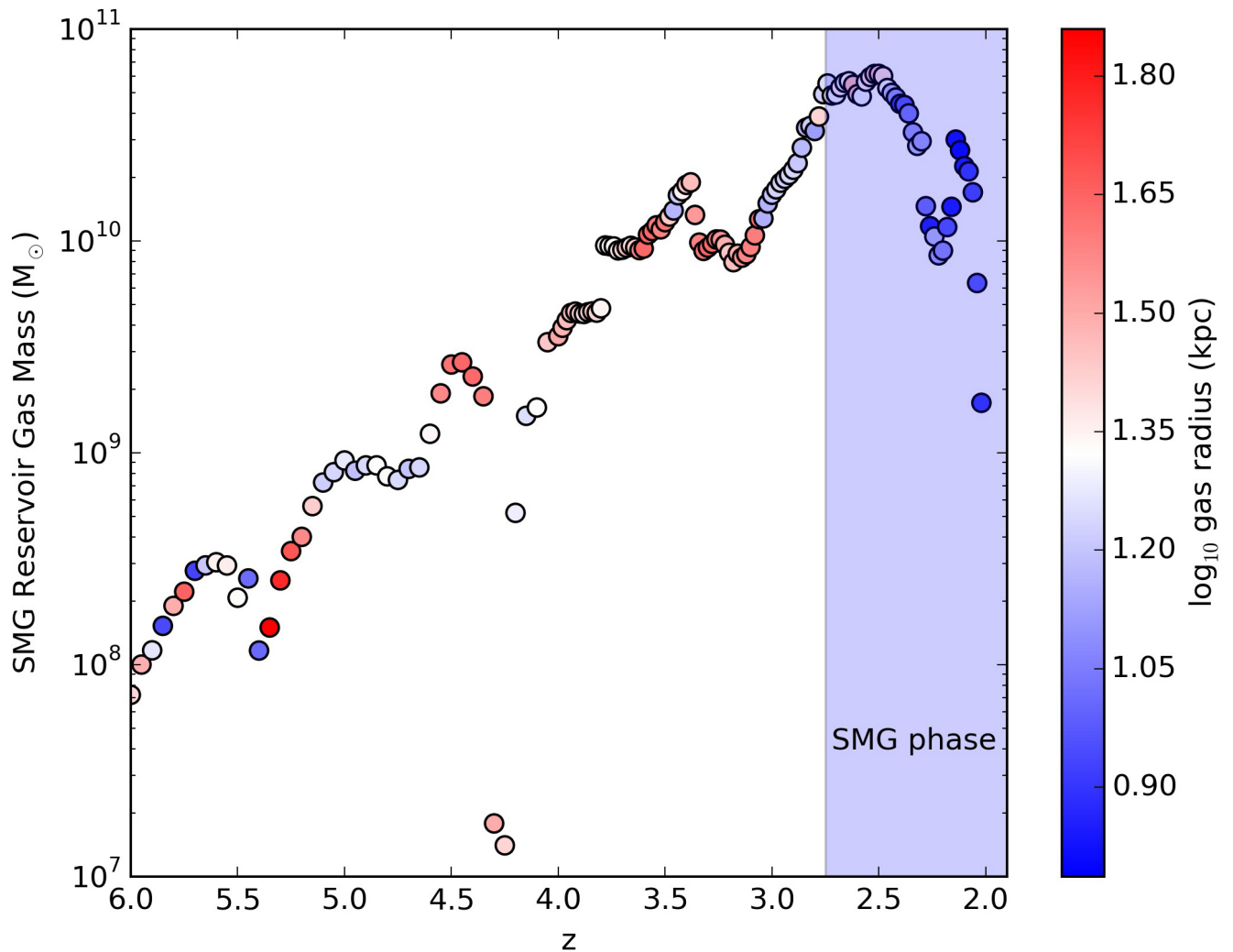
To fully capture the cosmic environment of SMGs during their formation, as well as their baryonic structure and morphology, cosmological hydrodynamic simulations are probably the best tool. Thus far, cosmological hydrodynamic simulations used to simulate SMGs have not employed direct radiative transfer models¹⁰. As such, inferring when a galaxy is an SMG in cosmological simulations has necessitated the use of parameterized emission models, such as assumed greybody emission laws⁹, or SFR thresholds¹⁰. The physical properties for SMGs derived from the most extensive of these studies¹⁰ (that is, M_* , M_{DM} and f_{gas}) are similar to the model presented here, although with roughly a factor of ~ 3 difference in SFR.

Code availability. We have made POWDERDAY available at <https://bitbucket.org/desika/powderday>, and GIZMO available at <https://bitbucket.org/phopkins/gizmo>.

31. Hopkins, P. F. A general class of Lagrangian smoothed particle hydrodynamics methods and implications for fluid mixing problems. *Mon. Not. R. Astron. Soc.* **428**, 2840–2856 (2013).
32. Agertz, O. *et al.* Fundamental differences between SPH and grid methods. *Mon. Not. R. Astron. Soc.* **380**, 963–978 (2007).

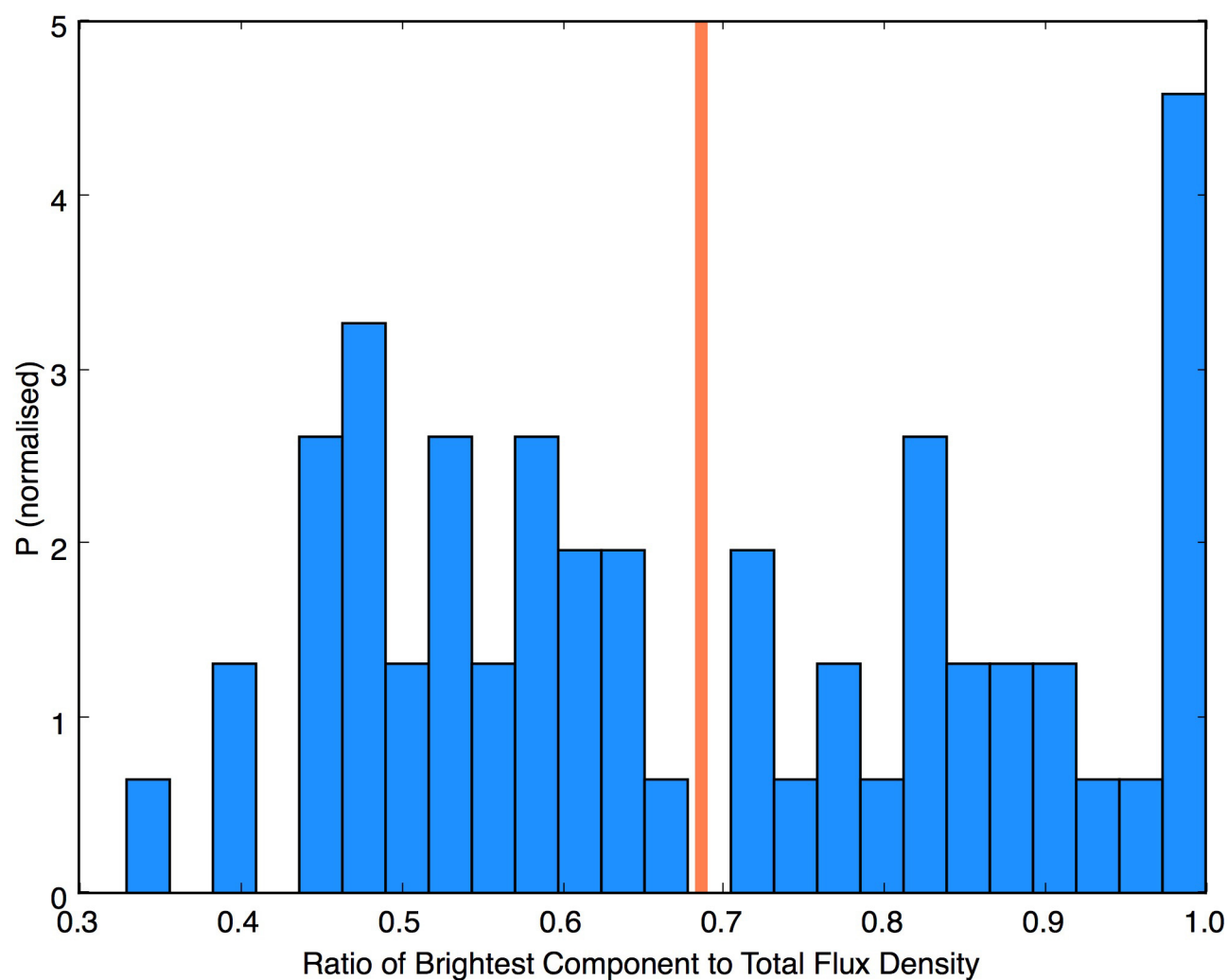
33. Sijacki, D., Vogelsberger, M., Kereš, D., Springel, V. & Hernquist, L. Moving mesh cosmology: the hydrodynamics of galaxy formation. *Mon. Not. R. Astron. Soc.* **424**, 2999–3027 (2012).
34. Hayward, C. C., Torrey, P., Springel, V., Hernquist, L. & Vogelsberger, M. Galaxy mergers on a moving mesh: a comparison with smoothed particle hydrodynamics. *Mon. Not. R. Astron. Soc.* **442**, 1996–2016 (2014).
35. Faucher-Giguère, C.-A. *et al.* Neutral hydrogen in galaxy haloes at the peak of the cosmic star formation history. *Mon. Not. R. Astron. Soc.* **449**, 987–1003 (2015).
36. Springel, V., Di Matteo, T. & Hernquist, L. Modelling feedback from stars and black holes in galaxy mergers. *Mon. Not. R. Astron. Soc.* **361**, 776–794 (2005).
37. Barnes, J. E. Gravitational softening as a smoothing operation. *Mon. Not. R. Astron. Soc.* **425**, 1104–1120 (2012).
38. Hahn, O. & Abel, T. Multi-scale initial conditions for cosmological simulations. *Mon. Not. R. Astron. Soc.* **415**, 2101–2121 (2011).
39. Hopkins, P. F., Quataert, E. & Murray, N. Self-regulated star formation in galaxies via momentum input from massive stars. *Mon. Not. R. Astron. Soc.* **417**, 950–973 (2011).
40. Hopkins, P. F., Quataert, E. & Murray, N. The structure of the interstellar medium of star-forming galaxies. *Mon. Not. R. Astron. Soc.* **421**, 3488–3521 (2012).
41. Hopkins, P. F., Narayanan, D., Murray, N. & Quataert, E. Dense molecular gas: a sensitive probe of stellar feedback models. *Mon. Not. R. Astron. Soc.* **433**, 69–77 (2013).
42. Hopkins, P. F. *et al.* Star formation in galaxy mergers with realistic models of stellar feedback and the interstellar medium. *Mon. Not. R. Astron. Soc.* **430**, 1901–1927 (2013).
43. Hopkins, P. F. *et al.* Resolving the generation of starburst winds in Galaxy mergers. *Mon. Not. R. Astron. Soc.* **433**, 78–97 (2013).
44. Narayanan, D. & Hopkins, P. F. Why is the Milky Way X-factor constant? *Mon. Not. R. Astron. Soc.* **433**, 1223–1229 (2013).
45. Katz, N., Weinberg, D. H. & Hernquist, L. Cosmological simulations with TreeSPH. *Astrophys. J. Suppl. Ser.* **105**, 19–35 (1996).
46. Ferland, G. J. *et al.* The 2013 Release of Cloudy. *Rev. Mex. Astron. Astrofis.* **49**, 137–163 (2013).
47. Krumholz, M. R., McKee, C. F. & Tumlinson, J. The atomic-to-molecular transition in galaxies. I. An analytic approximation for photodissociation fronts in finite clouds. *Astrophys. J.* **689**, 865–882 (2008).
48. Krumholz, M. R. & Gnedin, N. Y. A comparison of methods for determining the molecular content of model galaxies. *Astrophys. J.* **729**, 36 (2011).
49. Hopkins, P. F., Narayanan, D. & Murray, N. The meaning and consequences of star formation criteria in galaxy models with resolved stellar feedback. *Mon. Not. R. Astron. Soc.* **432**, 2647–2653 (2013).
50. Kroupa, P. The initial mass function of stars: evidence for uniformity in variable systems. *Science* **295**, 82–91 (2002).
51. Leitherer, C. *et al.* Starburst99: synthesis models for galaxies with active star formation. *Astrophys. J. Suppl. Ser.* **123**, 3–40 (1999).
52. Mannucci, F., Della Valle, M. & Panagia, N. Two populations of progenitors for Type Ia supernovae? *Mon. Not. R. Astron. Soc.* **370**, 773–783 (2006).
53. Bellovary, J. *et al.* The relative role of galaxy mergers and cosmic flows in feeding black holes. *Astrophys. J.* **779**, 136 (2013).
54. Anglés-Alcázar, D. *et al.* Torque-limited growth of massive black holes in galaxies across cosmic time. *Astrophys. J.* **800**, 127 (2015).
55. Hopkins, P. F., Kocevski, D. D. & Bundy, K. Do we expect most AGN to live in discs? *Mon. Not. R. Astron. Soc.* **445**, 823–834 (2014).
56. Kocevski, D. D. *et al.* CANDELS: constraining the AGN-merger connection with host morphologies at $z \sim 2$. *Astrophys. J.* **744**, 148 (2012).
57. Treister, E., Schawinski, K., Urry, C. M. & Simmons, B. D. Major galaxy mergers only trigger the most luminous active galactic nuclei. *Astrophys. J.* **758**, L39 (2012).
58. Governato, F. *et al.* The Local Group as a test of cosmological models. *New Astron. Rev.* **2**, 91–106 (1997).
59. Stadel, J. G. *Cosmological N-body Simulations and their Analysis*. Ph.D. thesis, Univ. Washington (2001).
60. Thompson, R. pyGadgetReader: GADGET snapshot reader for python. Astrophysics Source Code Library 1411.001 (2014).
61. Thompson, R. SPHGR: Smoothed-Particle Hydrodynamics Galaxy Reduction. Astrophysics Source Code Library 1502.012 (2015).
62. Turk, M. J. *et al.* yt: A multi-code analysis toolkit for astrophysical simulation data. *Astrophys. J. Suppl. Ser.* **192**, 9 (2011).
63. Conroy, C., Gunn, J. E. & White, M. The propagation of uncertainties in stellar population synthesis modeling. I. The relevance of uncertain aspects of stellar evolution and the initial mass function to the derived physical properties of galaxies. *Astrophys. J.* **699**, 486–506 (2009).
64. Conroy, C. & Gunn, J. E. The propagation of uncertainties in stellar population synthesis modeling. III. Model calibration, comparison, and evaluation. *Astrophys. J.* **712**, 833–857 (2010).
65. Robitaille, T. P. HYPERION: an open-source parallelized three-dimensional dust continuum radiative transfer code. *Astron. Astrophys.* **536**, A79 (2011).
66. Lucy, L. B. Computing radiative equilibria with Monte Carlo techniques. *Astron. Astrophys.* **344**, 282–288 (1999).
67. Weingartner, J. C. & Draine, B. T. Dust grain-size distributions and extinction in the Milky Way, Large Magellanic Cloud, and Small Magellanic Cloud. *Astrophys. J.* **548**, 296–309 (2001).
68. Robitaille, T. P. *et al.* A self-consistent model of Galactic stellar and dust infrared emission and the abundance of polycyclic aromatic hydrocarbons. *Astron. Astrophys.* **545**, A39 (2012).
69. Dwek, E. The evolution of the elemental abundances in the gas and dust phases of the galaxy. *Astrophys. J.* **501**, 643–665 (1998).

70. Vladilo, G. Dust and elemental abundances in damped Ly α absorbers. *Astrophys. J.* **493**, 583–594 (1998).
71. Watson, D. The Galactic dust-to-metals ratio and metallicity using gamma-ray bursts. *Astron. Astrophys.* **533**, A16 (2011).
72. Pascucci, I. *et al.* The 2D continuum radiative transfer problem. Benchmark results for disk configurations. *Astron. Astrophys.* **417**, 793–805 (2004).
73. Jonsson, P. SUNRISE: polychromatic dust radiative transfer in arbitrary geometries. *Mon. Not. R. Astron. Soc.* **372**, 2–20 (2006).
74. Jonsson, P., Groves, B. A. & Cox, T. J. High-resolution panchromatic spectral models of galaxies including photoionization and dust. *Mon. Not. R. Astron. Soc.* **186** (2010).
75. Torrey, P. *et al.* Synthetic galaxy images and spectra from the Illustris simulation. *Mon. Not. R. Astron. Soc.* **447**, 2753–2771 (2015).
76. Ade, P. A. R. *et al.* Planck 2013 results. XVI. Cosmological parameters. *Astron. Astrophys.* **571**, A16 (2014).
77. González, J. E., Lacey, C. G., Baugh, C. M. & Frenk, C. S. The role of submillimetre galaxies in hierarchical galaxy formation. *Mon. Not. R. Astron. Soc.* **413**, 749–762 (2011).
78. Hainline, L. J. *et al.* The stellar mass content of submillimeter-selected galaxies. *Astrophys. J.* **740**, 96 (2011).
79. Michałowski, M. J. *et al.* Determining the stellar masses of submillimetre galaxies: the critical importance of star formation histories. *Astron. Astrophys.* **571**, A75 (2014).
80. Bothwell, M. S. *et al.* A survey of molecular gas in luminous sub-millimetre galaxies. *Mon. Not. R. Astron. Soc.* **429**, 3047–3067 (2013).
81. Narayanan, D., Bothwell, M. & Davé, R. Galaxy gas fractions at high redshift: the tension between observations and cosmological simulations. *Mon. Not. R. Astron. Soc.* **426**, 1178–1184 (2012).
82. Tacconi, L. J. *et al.* Submillimeter galaxies at $z \sim 2$: evidence for major mergers and constraints on lifetimes, IMF, and CO-H₂ conversion factor. *Astrophys. J.* **680**, 246–262 (2008).
83. Chakrabarti, S., Fenner, Y., Cox, T. J., Hernquist, L. & Whitney, B. A. An evolutionary model for submillimeter galaxies. *Astrophys. J.* **688**, 972–989 (2008).
84. Narayanan, D., Cox, T. J., Hayward, C. C., Younger, J. D. & Hernquist, L. The star-forming molecular gas in high-redshift submillimetre galaxies. *Mon. Not. R. Astron. Soc.* **400**, 1919–1935 (2009).
85. Narayanan, D. *et al.* The formation of high-redshift submillimetre galaxies. *Mon. Not. R. Astron. Soc.* **401**, 1613–1619 (2010).
86. Hayward, C. C. *et al.* What does a submillimeter galaxy selection actually select? The dependence of submillimeter flux density on star formation rate and dust mass. *Astrophys. J.* **743**, 159 (2011).
87. Narayanan, D., Krumholz, M. R., Ostriker, E. C. & Hernquist, L. A general model for the CO-H₂ conversion factor in galaxies with applications to the star formation law. *Mon. Not. R. Astron. Soc.* **421**, 3127–3146 (2012).

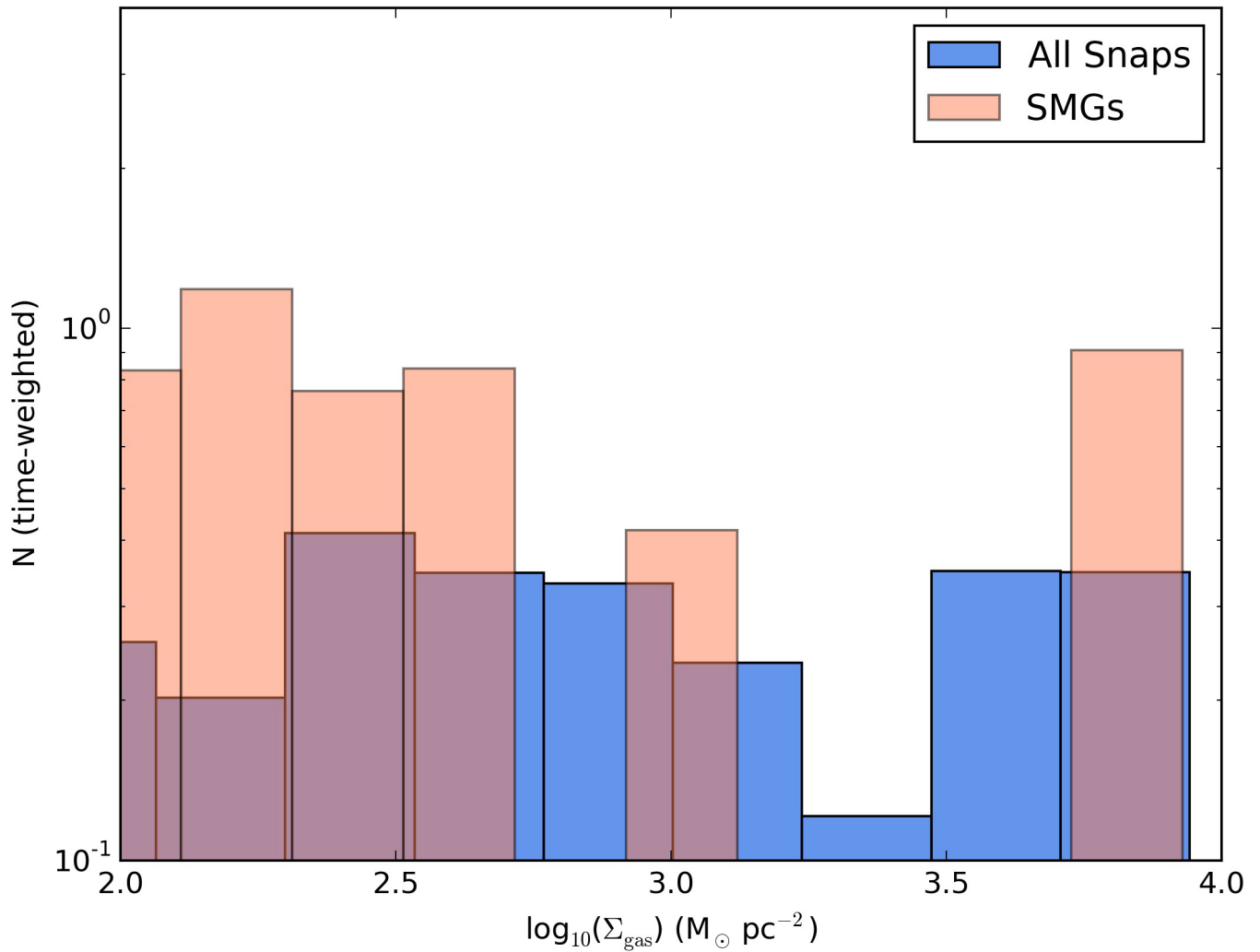


Extended Data Figure 1 | Mass of reservoir gas in the central galaxy that will be consumed during SMG starburst as a function of z . The colour scale denotes the median scale height from the galaxy centre of mass. The gas mass consumed during the starburst is calculated by tracking the evolution of gas

particles that turn into stars during the SMG phase ($z \approx 2-2.7$), and is only measured for the central galaxy itself (that is, gas ejected into the halo is not included). The SMG gas reservoir follows a cycle of being pushed outward followed by re-accretion.

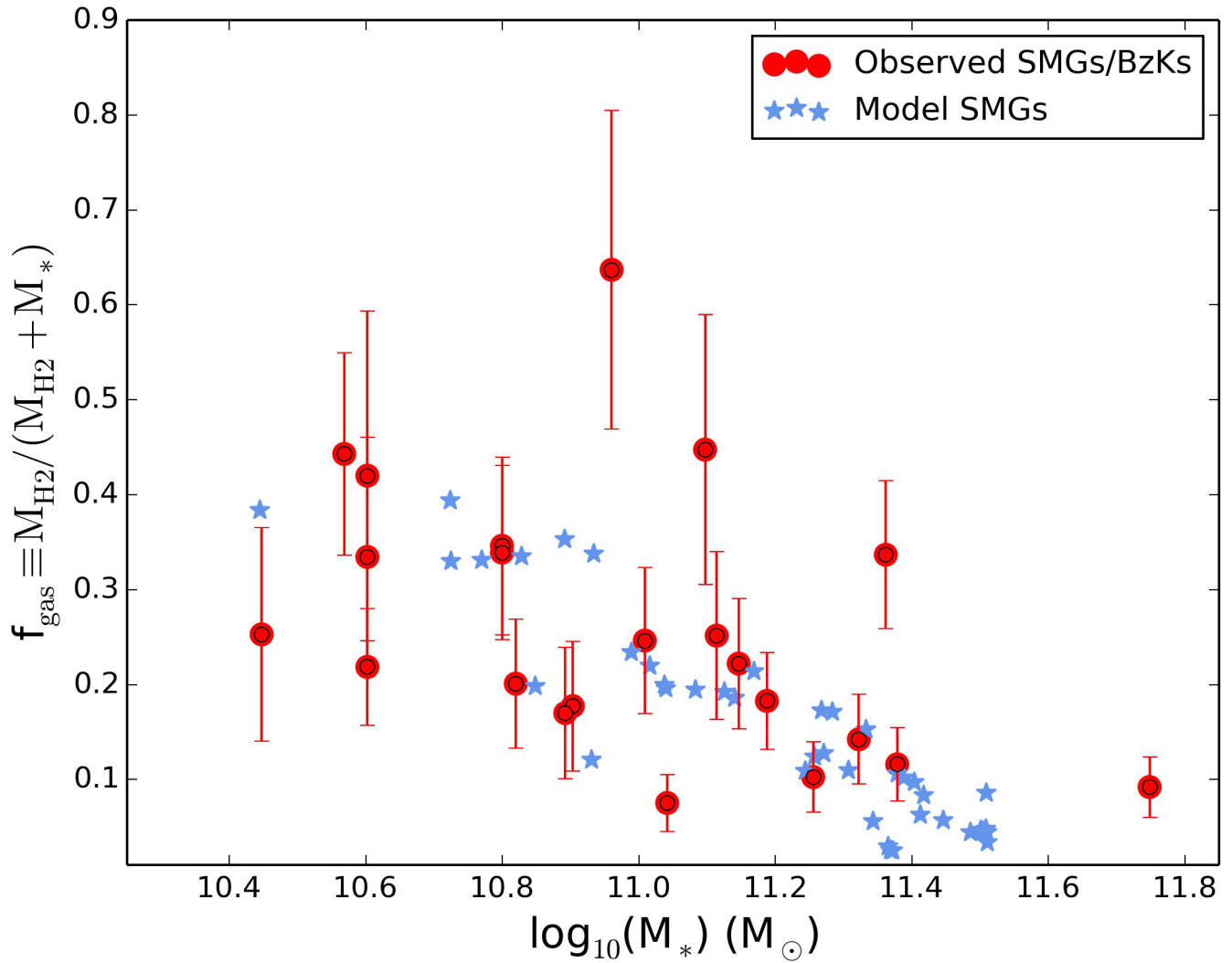


Extended Data Figure 2 | Distribution of flux density ratio of brightest component in submillimetre-luminous region to total flux density. The average is shown with the vertical line. Submillimetre-luminous regions often break up into multiples. The region is generally dominated by one component, although smaller subhaloes can contribute on average ~30% of the observed flux density. The normalization of the ordinate, P , is arbitrary.



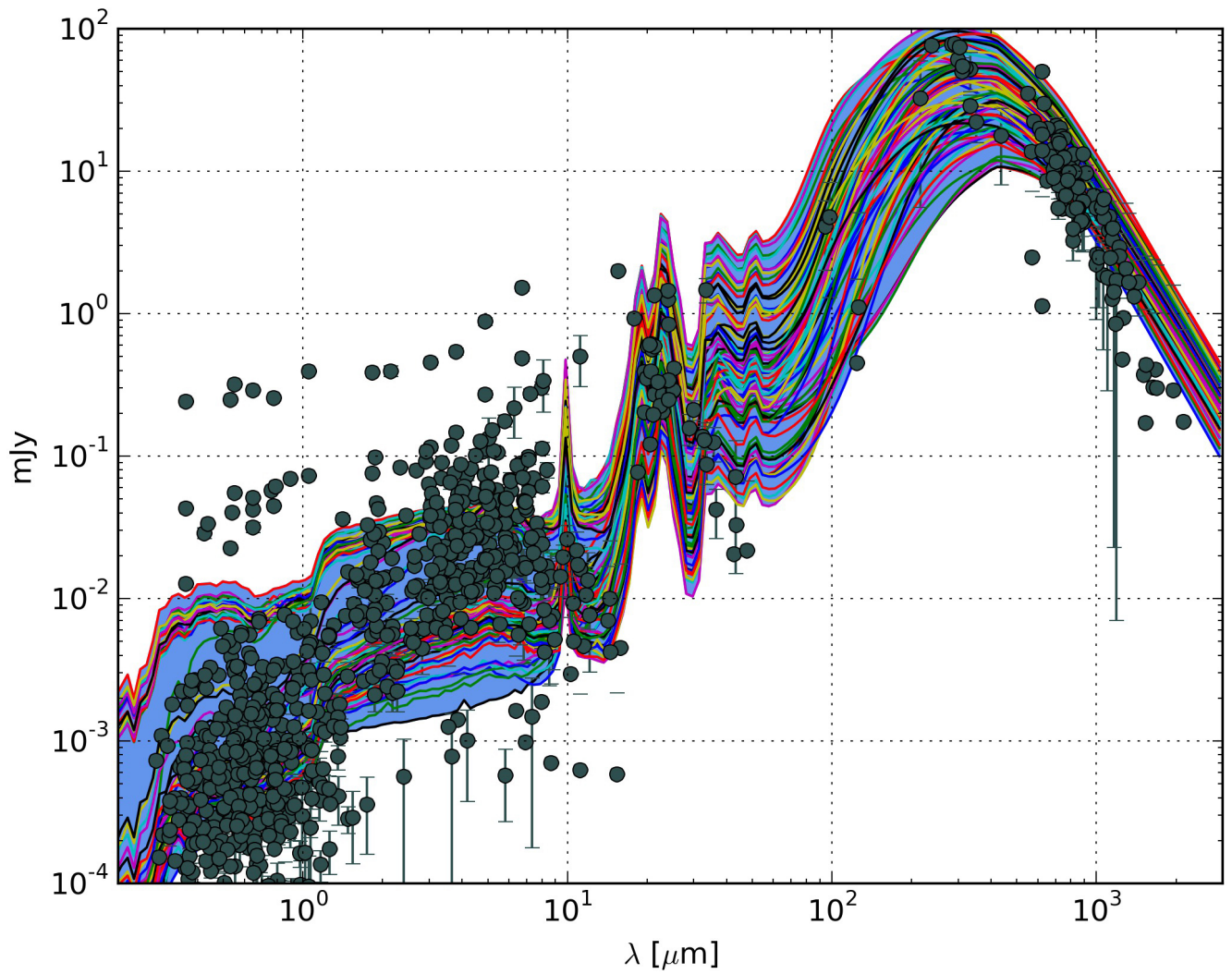
Extended Data Figure 3 | Gas surface density for the central submillimetre galaxy. The blue histogram shows the distribution of gas surface densities (Σ_{gas}) during all phases (that is, all snapshots, Snaps), while the pink histogram shows the same for the submillimetre-luminous phase. The ordinate (N) is weighted by the time a galaxy spends in a given gas surface density bin, and the

normalization is arbitrary. We predict that the submillimetre-luminous phases do not have dramatically different surface density distributions compared to the non-submillimetre-luminous phases. This prediction might have been tentatively observed^{1,87}.



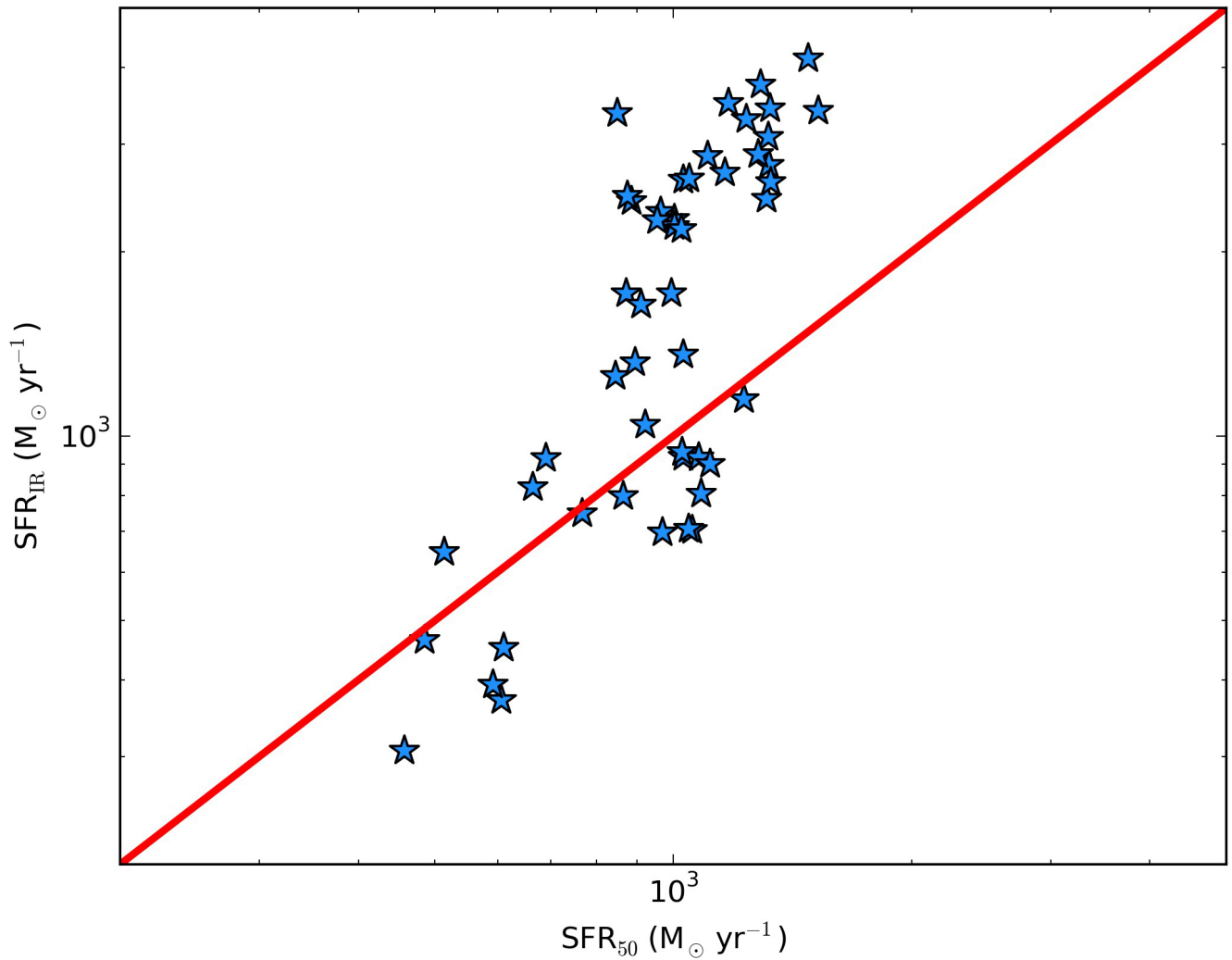
Extended Data Figure 4 | Molecular gas fraction as a function of galaxy stellar mass. Blue stars show individual snapshots of the central submillimetre galaxy, while red circles with error bars (1σ) show observations of BzK galaxies and SMGs with direct CO($J = 1-0$) measurements (to avoid complications in converting from higher-lying CO rotational lines to the

ground state for a mass conversion). (BzK galaxies are those that have been selected on the basis of their B, z and K band luminosities.) Both the observations and our model show a declining molecular gas fraction (f_{gas}) with increasing galaxy mass (M_*), with a typical range of $f_{\text{gas}} = 0.1-0.4$ for galaxies of SMG mass.



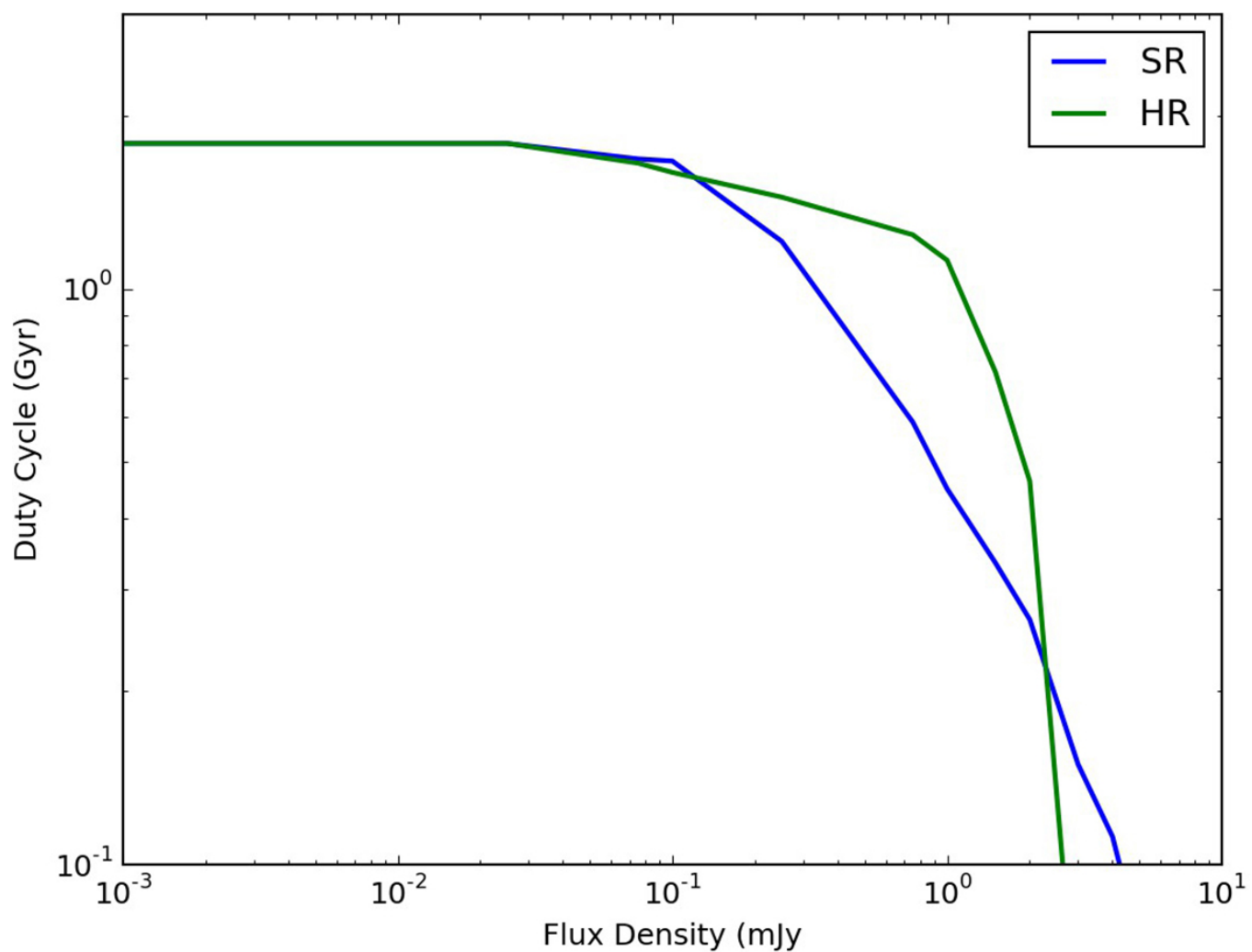
Extended Data Figure 5 | Predicted spectral energy distribution (SED) for the central submillimetre galaxy. The ordinate shows the flux density in mJy, while the abscissa shows the wavelength in μm . The blue shaded region shows the range of SEDs for all simulation snapshots that satisfy the fiducial $F_{850\mu\text{m}} > 5$ mJy submillimetre galaxy selection criteria, while the dark grey

points with error bars (1σ) are a compilation of observed data. The individual coloured lines show the SEDs for individual submillimetre-luminous snapshots. The data and models are redshifted to a common redshift $z = 2$. The model and data compare well, and the model suggests a diverse range of SMG SEDs.



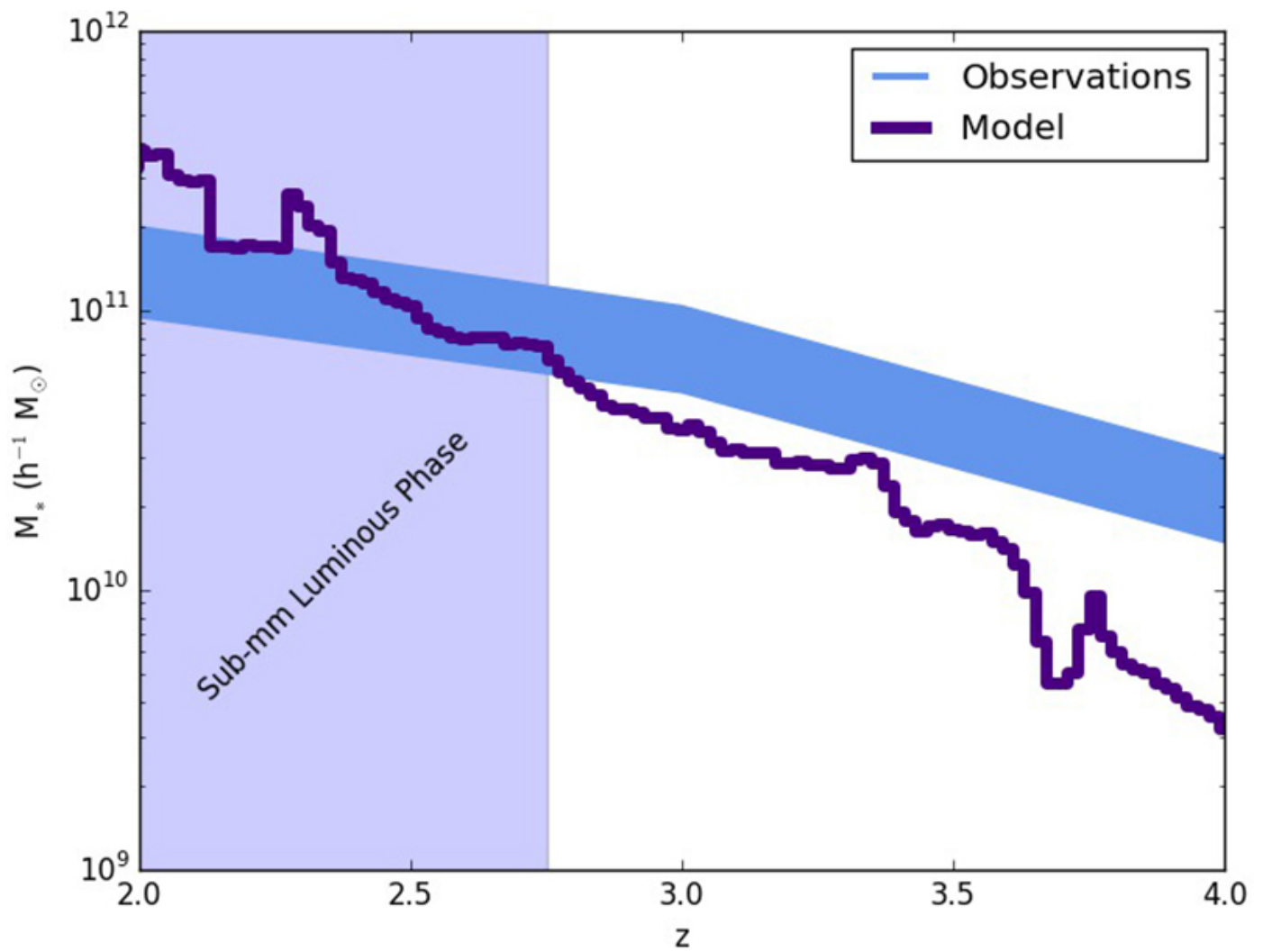
Extended Data Figure 6 | Overestimate of the SFR of high- z SMGs. The ordinate denotes the SFR as determined from the infrared SED (SFR_{IR})²⁵, while the abscissa shows the SFR averaged over the last 50 Myr in the simulations

(SFR_{50}). Up to an SFR of $\sim 800 M_{\odot} \text{ yr}^{-1}$ the two correspond well. At higher SFRs, however, there is a dramatic departure owing to substantial contribution to the infrared luminosity by older stars.



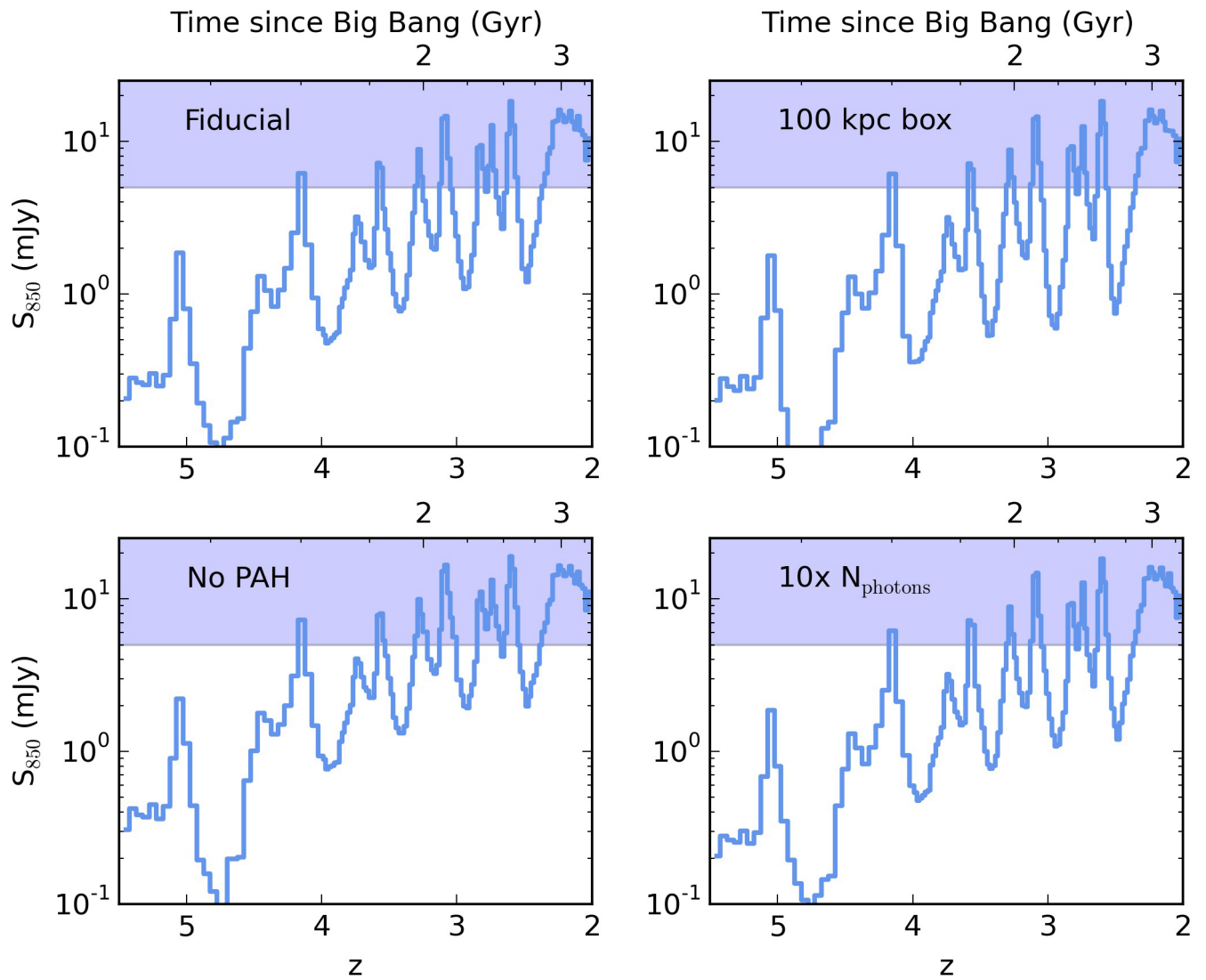
Extended Data Figure 7 | Resolution tests for hydrodynamic zoom simulations. Lines show the 850 μm duty cycle above a given flux density as a function of flux density for our resolution test models presented in Methods. SR

denotes our standard resolution (the resolution of our main model) while HR is a one-level-higher refinement.



Extended Data Figure 8 | Stellar mass–redshift relation for the model galaxy. The purple line shows model results, while the dark-blue filled region shows observational constraints from an abundance matching assumption²². The model and observations are in reasonable agreement, especially during

the submillimetre-luminous phase (vertical shaded region). At late times, the stellar mass of the galaxy is a factor of ~ 2 higher than the median observed galaxy.



Extended Data Figure 9 | Tests of parameter choices for radiative transfer calculations. The simulated galaxy for these tests is our lowest resolution cosmological simulation (m13m14). Each panel shows the 850 μm flux density light curve of the tested model, with time noted on the abscissa (redshift on the bottom, time since the Big Bang on the top). In all panels, the shaded region

denotes $S_{850} \geq 5$ mJy, which is the canonical selection criteria for SMGs. Top left, our fiducial set of parameters; top right, simulation with a 100 kpc (on a side) emission region instead of 200 kpc; bottom left, simulation with our model for PAHs turned off; bottom right, fiducial simulation run with ten times the number of photons.

Extended Data Table 1 | Summary of model galaxies

Model Name	Model Purpose	M_* (z=2) M_\odot	M_{halo} (z=2) M_\odot	m_b M_\odot	m_{DM} M_\odot	ϵ_b pc	ϵ_{DM} pc	Final Redshift
B100	Main Model	4×10^{11}	3×10^{13}	2.7×10^5	1.3×10^6	9	142	2
TL37 SR	Resolution Test	8×10^{10}	7×10^{12}	2.7×10^5	1.3×10^6	9	142	2
TL37 HR	Resolution Test	5×10^{10}	7×10^{12}	3.3×10^4	1.7×10^5	9	142	2
m13m14	RT Parameter Survey	3×10^{11}	7×10^{12}	4.4×10^6	2.3×10^7	70	700	0.2

M_* and M_{halo} refer to the stellar and halo mass at $z = 2$; ϵ_b and ϵ_{DM} refer to the minimum force softening lengths for baryons and dark matter particles, respectively; and m_b and m_{DM} refer to the baryonic and dark matter particle masses, respectively. For model m13m14, RT stands for radiative transfer.

The diurnal cycle of water ice on comet 67P/Churyumov–Gerasimenko

M. C. De Sanctis¹, F. Capaccioni¹, M. Ciarniello¹, G. Filacchione¹, M. Formisano¹, S. Mottola², A. Raponi¹, F. Tosi¹, D. Bockelée-Morvan³, S. Erard³, C. Leyrat³, B. Schmitt⁴, E. Ammannito^{1,5}, G. Arnold², M. A. Barucci³, M. Combi⁶, M. T. Capria¹, P. Cerroni¹, W.-H. Ip⁷, E. Kuehrt², T. B. McCord⁸, E. Palomba¹, P. Beck⁴, E. Quirico⁴ & the VIRTIS Team*

Observations of cometary nuclei have revealed a very limited amount of surface water ice^{1–7}, which is insufficient to explain the observed water outgassing. This was clearly demonstrated on comet 9P/Tempel 1, where the dust jets (driven by volatiles) were only partially correlated with the exposed ice regions⁸. The observations^{6,7} of 67P/Churyumov–Gerasimenko have revealed that activity has a diurnal variation in intensity arising from changing insolation conditions. It was previously concluded that water vapour was generated in ice-rich subsurface layers with a transport mechanism linked to solar illumination^{1–3,5}, but that has not hitherto been observed. Periodic condensations of water vapour very close to, or on, the surface were suggested^{3,9} to explain short-lived outbursts seen near sunrise on comet 9P/Tempel 1. Here we report observations of water ice on the surface of comet 67P/Churyumov–Gerasimenko, appearing and disappearing in a cyclic pattern that follows local illumination conditions, providing a source of localized activity. This water cycle appears to be an important process in the evolution of the comet, leading to cyclical modification of the relative abundance of water ice on its surface.

The Visible Infrared and Thermal Imaging Spectrometer VIRTIS¹⁰ has collected data of high spatial (7–25 m per pixel) and spectral resolution since the Rosetta spacecraft approached the nucleus of comet 67P/Churyumov–Gerasimenko in August 2014. The reflectance spectra, taken in different areas over the illuminated regions of the comet's nucleus, show a broad absorption band at 2.8–3.6 μm ,

attributed to organic compounds. The absence of pure water ice absorption bands indicates an upper limit of about 1% (by volume) of water ice, in very limited surface regions, at VIRTIS resolution⁴.

Figure 1 shows a small region of the ‘neck’ (longitude $325^\circ \pm 4^\circ$ E, latitude $31^\circ \pm 5^\circ$ N, called Hapi) of the comet, located between the small and large lobes of the nucleus, observed at different rotational phases after one or more comet rotations. During each rotation, this region moves into the shadows projected by the head (the smaller lobe) bulge. VIRTIS observes variations in the absorption band near 3 μm as this region moves out of the shadow and becomes illuminated (Fig. 2a). In this case we observe a clear alteration of the organic compounds band, with a broadening, a shift towards shorter wavelengths and a strong increase in depth with decreasing illumination. The band centre shifts from $3.22 \pm 0.05 \mu\text{m}$ to $3.12 \pm 0.05 \mu\text{m}$, its short-wavelength shoulder shifts from $2.82 \pm 0.05 \mu\text{m}$ to $2.71 \pm 0.05 \mu\text{m}$, and the band depth relative to the continuum increases from 0.20 ± 0.03 to 0.60 ± 0.03 . We also observe a flattening of the continuum slope and a reduced thermal emission as the 3- μm band depth increases.

The stronger 3- μm band and the flattening of the continuum suggest exposed water ice, in addition to the organic material ubiquitously present on the comet's surface⁴. The spectral ratio (Fig. 2) between the spectrum of a pixel close to the shadow and the spectrum of an illuminated pixel shows the presence of a strong ice band extending from 2.7 μm to 3.6 μm , while the other water ice bands at 1.5 μm and 2.0 μm are not detected¹¹ (see Extended Data Fig. 1). The 3- μm water band is

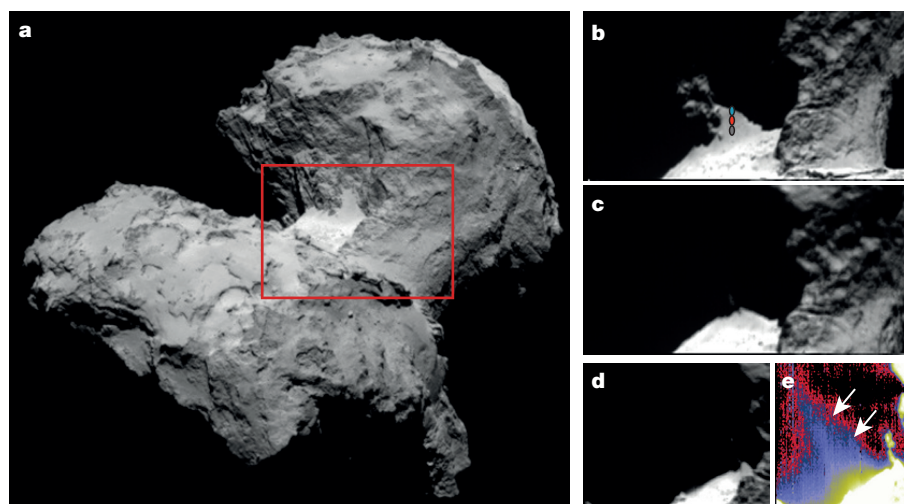


Figure 1 | Images of the ice-rich area. **a**, Rosetta Optical Navigation Camera (OPNAV) context image of the region under study (red box); **b–e**, VIRTIS image at 0.7 μm of the region in the red box of **a**. The data in **b**, **c** and **d** were acquired on 12, 13 and 14 September 2014, respectively. The VIRTIS data in **b** and **c** are separated by ~ 12 h, corresponding to ~ 1 comet rotation, while the data in **c** and **d** are separated by 37.3 h, corresponding to ~ 3 comet rotations. The coloured dots in **b** indicate the zones from which the spectra in Fig. 2 are taken. Panel **e** is the same as **d**, but stretched to see the jet (white arrows).

¹Istituto di Astrofisica e Planetologia Spaziali – INAF, via del fosso del cavaliere 100, 00133 Rome, Italy. ²Institute for Planetary Research, DLR, Rutherfordstraße 2, 12489 Berlin, Germany. ³LESIA-Observatoire de Paris, CNRS, Université Pierre et Marie Curie, Université Paris Diderot, 5 place Jules Janssen, 92195 Meudon, France. ⁴Université Grenoble Alpes – CNRS Institut de Planetologie et Astrophysique de Grenoble, Batiment D de Physique, BP 53, 38041 Grenoble Cedex 9, France. ⁵University of California, Los Angeles, California 90095, USA. ⁶Department of Atmospheric, Oceanic and Space Sciences, University of Michigan, 2455 Hayward Street, Ann Arbor, Michigan 48109, USA. ⁷National Central University, No. 300, Zhongda Road, Zhongli District, Taoyuan City, 32001 Taipei, Taiwan. ⁸Bear Fight Institute, 22 Fiddler's Road, Box 667, Winthrop, Washington 98862, USA.

*A list of participants and their affiliations is given at the end of this paper.

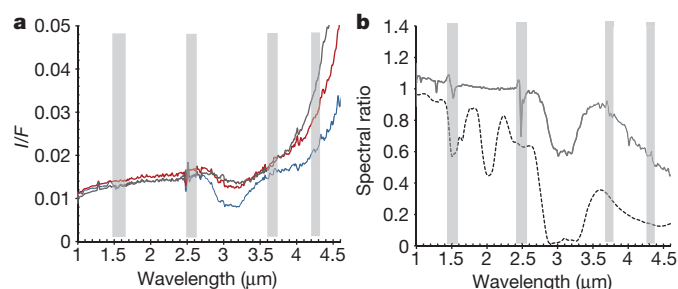


Figure 2 | Spectra of the ice-rich areas. **a**, Spectra from Fig. 1b going from illuminated pixels to shadow. Black, red and cyan spots in Fig. 1b correspond to black, red and cyan spectra, respectively, taken at steps of 1 pixel. At wavelengths >3.5 – $3.7\ \mu\text{m}$, the spectra show smaller thermal emission for the pixels closer to the shadow line. The retrieved temperatures are $175 \pm 8\ \text{K}$, $184 \pm 5\ \text{K}$ and $195 \pm 4\ \text{K}$ for the cyan, red and black spots, respectively. **b**, Spectral ratio of the cyan and black spectrum of **a** (solid line) and a synthetic spectrum of water¹¹ (grain size of $10\ \mu\text{m}$, dashed line). Instrument filters are indicated by grey bars.

clearly present in all the pixels located at the border of the shadowed region. The same region has been observed again one and four nucleus revolutions¹² later (Fig. 1b–d) under slightly different illumination conditions, as shown by the shadows which cover different areas. Nevertheless, in each observation, the presence and change in the

water absorption band depth is controlled by the shadow location and not by the specific surface region. VIRTIS observes areas in which the spectra display progressive $3\text{-}\mu\text{m}$ band weakening as the region moves into greater illumination. Shadowed areas are slightly different from the illuminated regions in the VIRTIS observations (Fig. 1b–d), and ice-rich and ice-free pixels change according to their distance from the shadow: ice-free pixels in the first observation (Fig. 1b) show the ice signature ('ice-rich') on the following observation (Fig. 1c, d) where they are now closer to the shadow.

Using optical constants^{13–16} and scattering theory¹⁷, we modelled the spectra as an intimate mixture of ice and a dark non-ice component (Methods, Extended Data Figs 1 and 2) and derived water-ice abundance maps (Fig. 3). The fit of the spectra requires a relative abundance of up to 10–15% of water ice intimately mixed with the non-ice component, as shown in the ice distribution maps (Fig. 3a). The maps indicate that the maximum quantity of ice is found very close to the shadows in all the observations, even if the pixels close to the shadows are at a different location on the comet surface.

The spectra also show differences at wavelengths longer than $3.6\ \mu\text{m}$, owing to variations of the thermal emission as the region moves into reduced solar illumination, with a correlation of a stronger $3\text{-}\mu\text{m}$ band with a weaker thermal emission, thus with lower temperatures. The nucleus surface temperatures were retrieved from the long-wavelength portion ($>4.5\ \mu\text{m}$) of the VIRTIS spectra¹⁸. The temperatures retrieved for the pixels near the shadows are at most $160\ \text{K}$ with an uncertainty of

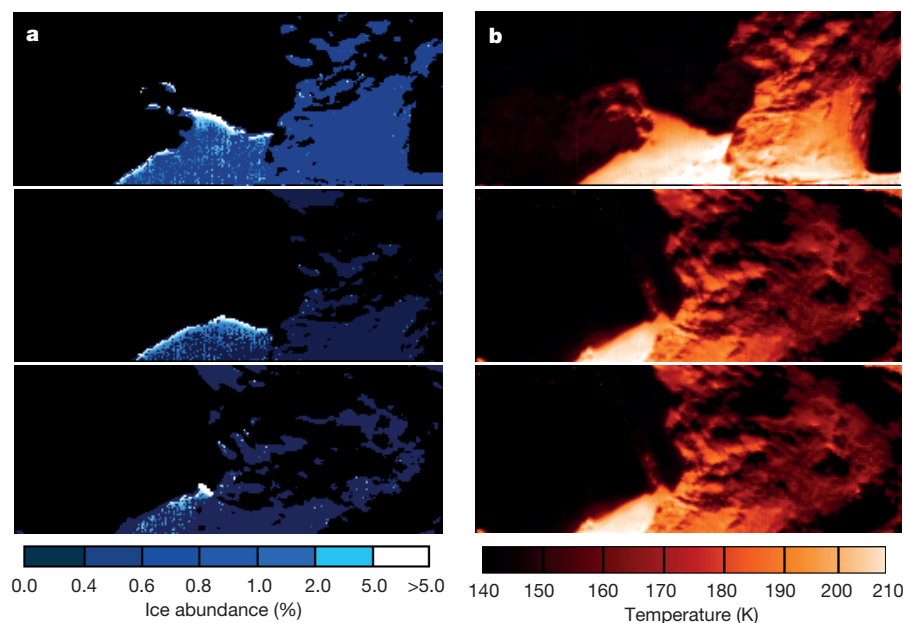
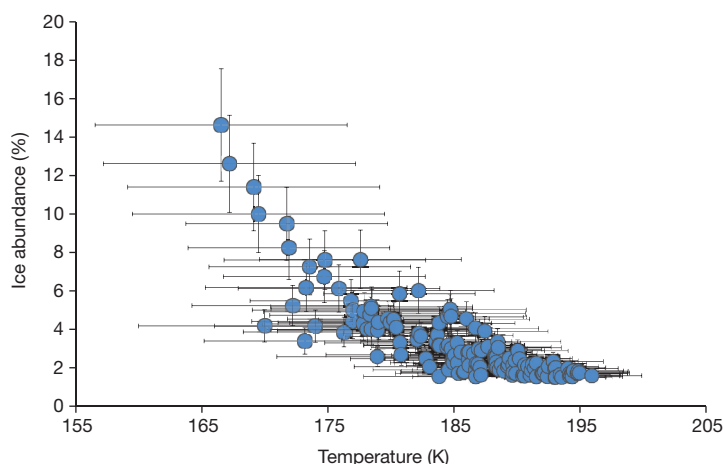


Figure 3 | Ice and temperature maps. **a**, Ice maps: ice abundance (by volume) as shown by the colour scale under. The black pixels are those in shadow or corresponding to sky. Isolated bright pixels are due to instrumental artefacts. **b**, Temperature images. White colours correspond to high temperatures while red-brown colours correspond to low temperatures (see colour scale under). **c**, Scatter plot of ice abundance versus temperature. The data points are extracted from **a** and refer to the region of the neck near the shadow. The error associated with the ice amount is 20%, and errors on the temperatures are $\sim 3\%$ above $170\ \text{K}$ and $\sim 5\%$ below $170\ \text{K}$.



± 10 K, while the temperature of the well illuminated areas are up to 210 K with an uncertainty of ± 2 K (Fig. 3b). The pixels showing the strongest 3- μm absorption have temperatures in the range 160–180 K and a derived ice content varying in a range between $5 \pm 1\%$ and $14 \pm 3\%$ (Fig. 3). The measured temperatures for these pixels are consistent with the presence of water ice in the nucleus outer layers. Water ice sublimation pressure¹⁹ varies by three orders of magnitude from temperatures in the shadowed areas (< 1 nbar), to temperatures in the illuminated areas, where the activity occurs. A clear anti-correlation of ice abundance with temperatures is seen (Fig. 3c) in the regions near the shadows of Fig. 1b, indicating that the temperature is cold enough to maintain ice at, or near, the surface in the shadows.

Most of the activity observed in the period August–September 2014 is from the comet neck^{6,7}, where the region previously examined is located (Fig. 1a). Activity from this specific region is seen in all the data presented here, especially in the data acquired last (Fig. 1e), indicating that a source of water exists to sustain the flux. The amount of water flux coming from the superficial ice documented by VIRTIS represents $\sim 3\%$ of the total water flux measured by MIRO⁶ (see Methods for calculations). From this, we can extrapolate that a much larger area is affected by the same mechanism (although this larger area is not directly observed) and is thus contributing to a larger amount of total water flux. Indeed, all the Hapi region is subjected to this diurnal shadowing effect which can lead to the outgassing over much larger areas. It must be mentioned that the contribution to the total outgassing from these surface layers sources is limited in time. This is demonstrated by the progressive decrease of the abundance of deposited water ice in pixels exposed for a longer time to the solar illumination (see Fig. 3). In the case studied here, the presence of surface ice close to the location of the jets indicates that the outgassing source is likely to be in the uppermost layers of the surface.

The above description of VIRTIS observations of ice sublimating in this neck region (Hapi) when an area emerges from shadow, and the progressive decrease of the ice abundance away from the shadow, clearly indicates that a cyclic sublimation–condensation process is at work during each comet rotation.

Two possible mechanisms for the cyclic condensation of water on unilluminated areas can be considered: (1) the condensation of water vapour present in the coma, and/or redeposition of icy grains, on cold areas on the nucleus^{20,21}, or (2) the direct condensation of gas sublimating from the subsurface under appropriate thermodynamic conditions^{9,22–24}. The first case could indeed occur in the region of the neck where, because of the large concavity, sublimated molecules from an illuminated region could impact and condense on nearby non-illuminated areas. However, this mechanism seems to be more efficient at small heliocentric distance when gas production rate is high enough to enhance the re-deposited flux significantly²⁰. The second mechanism has been already suggested to explain the outbursts observed by the Deep Impact mission on comet 9P/Tempel 1 that appear to occur near sunrise on a particular area⁹, while extensive subsurface sources were invoked to explain the overall ambient outgassing as the observed area of exposed pure ice has a too-limited extent². In addition, during the fly-by of 103P/Hartley 2, the DIXI mission has revealed surface ice notably only along the morning terminator, suggesting diurnal effects²⁵.

The VIRTIS observations are now able to demonstrate that a mechanism similar to (2) above is at work; our thermo-physical model of the nucleus, along with previous literature^{9,22,24}, enables the diurnal cycle of water to be quantified. When the surface is illuminated, water ice sublimates mainly from the uppermost surface layers (Fig. 4 and Extended Data Figs 3 and 4). When the surface goes into shadow (or into the night side), a temperature inversion occurs between the now colder surface layer and the interior layers which maintain a higher temperature for a longer time; the magnitude of this process is controlled by the duration of the shadow/night period and by the thermal inertia of the material and extends to a depth defined by the thermal skin depth. In the present case, the thermal inertia is constrained by

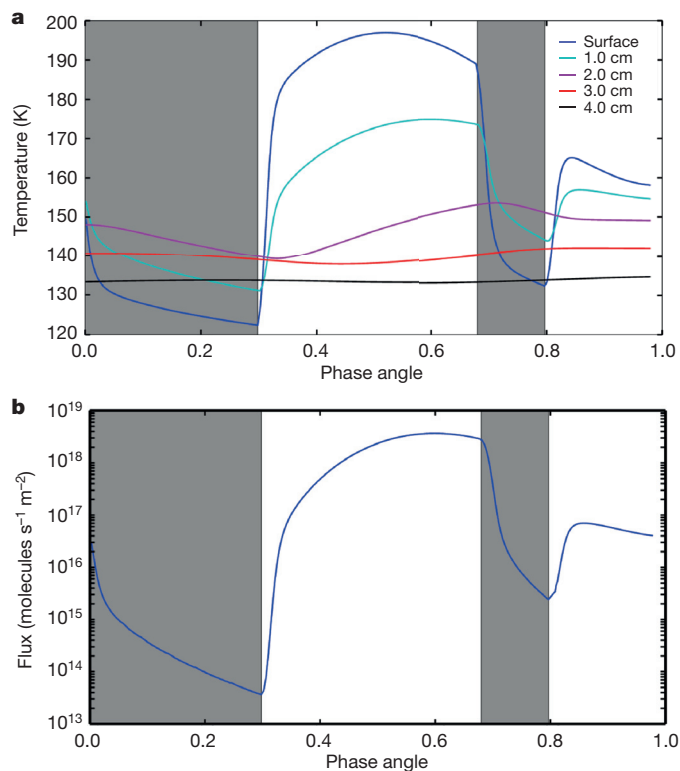


Figure 4 | Temperature and water vapour versus time. **a**, Simulated temperature at different depths from the nucleus surface to the interior (see key at right), over one rotational period, for a small region of the neck near the shadows. **b**, Simulated flux of water vapour versus time over one rotational period. The simulation has been done using the ‘Rome’ model^{22,24} (Methods) and the parameters reported in Extended Data Table 2. The grey bars correspond to unilluminated periods of time due to the combination of comet rotation and shadows.

independent measurements^{4,6} and gives rise to a thermal skin depth of the order of few centimetres (Fig. 4 and Extended Data Fig. 4).

Thus, within the few centimetres affected by the heat exchange, water vapour still produced by subsurface sublimation in the warmer subsurface layers flows through the pores and could re-condense if the thermo-physical conditions of the colder upper surface layer allow it. By this mechanism the surface layer becomes enriched in water ice. The water ice in the uppermost surface layers will be stable until a new cycle of solar illumination starts which will increase the surface temperature and thus trigger again the outgassing of water from the comet (Extended Data Fig. 3).

We thus suggest that the cyclic sublimation–condensation of ice triggered by varying illumination conditions is a general process acting on cometary nuclei. This process implies the cyclical modification of the relative abundance of water ice on the surface of the comet, contributing to the local water activity. This mechanism could lead to differential erosion of the surface, producing morphological differences or enhancing prior existing inhomogeneities. This mechanism could also contribute to activity arising from the pits on comet 67P/Churyumov–Gerasimenko²⁶, being relevant to the shadowing effects in the pits.

Moreover, the surface erosion of 67P/Churyumov–Gerasimenko could keep the water ice close to the surface, thus avoiding fading of activity with time. Finally, as this comet moves towards its perihelion, which will be reached in August 2015, the solar radiation will increase with the inverse square of the heliocentric distance, and thus this process of water sublimation–condensation will become progressively more energetic, possibly leading to the formation of outbursts events like those observed on the nucleus of 9P/Tempel 1.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 25 April; accepted 19 June 2015.

1. Soderblom, L. A. *et al.* Observations of Comet 19P/Borrelly by the miniature integrated camera and spectrometer aboard Deep Space 1. *Science* **296**, 1087–1091 (2002).
2. Sunshine, J. M. *et al.* Exposed water ice deposits on the surface of comet 9P/Tempel 1. *Science* **311**, 1453–1455 (2006).
3. A'Hearn, M. F. *et al.* EPOXI at comet Hartley 2. *Science* **332**, 1396–1400 (2011).
4. Capaccioni, F. *et al.* The organic-rich surface of comet 67P/Churyumov-Gerasimenko as seen by VIRTIS/Rosetta. *Science* **347**, <http://dx.doi.org/10.1126/science.aaa0628> (2015).
5. Feaga, L. M. *et al.* Asymmetries in the distribution of H₂O and CO₂ in the inner coma of Comet 9P/Tempel 1 as observed by Deep Impact. *Icarus* **190**, 345–356 (2007).
6. Gulkis, S. *et al.* Subsurface properties and early activity of comet 67P/Churyumov-Gerasimenko. *Science* **347**, <http://dx.doi.org/10.1126/science.aaa0709> (2015).
7. Sierk, H. *et al.* On the nucleus structure and activity of comet 67P/Churyumov-Gerasimenko. *Science* **347**, <http://dx.doi.org/10.1126/science.aaa1044> (2015).
8. Farnham, T. L. *et al.* Dust coma morphology in the Deep Impact images of Comet 9P/Tempel 1. *Icarus* **191**, 146–160 (2007).
9. Prialnik, D., A'Hearn, M. F. & Meech, K. J. A mechanism for short-lived cometary outbursts at sunrise as observed by Deep Impact on 9P/Tempel 1. *Mon. Not. R. Astron. Soc.* **388**, L20–L23 (2008).
10. Coradini, A. *et al.* An imaging spectrometer for the Rosetta mission. *Space Sci. Rev.* **128**, 529–559 (2007).
11. Filacchione, G. *et al.* Saturn's icy satellites and rings investigated by Cassini-VIMS: III — Radial compositional variability. *Icarus* **220**, 1064–1096 (2012).
12. Mottola, S. *et al.* The rotation state of 67P/Churyumov-Gerasimenko from approach observations with the OSIRIS cameras on Rosetta. *Astron. Astrophys.* **569**, L2 (2014).
13. Warren, S. G. Optical constants of ice from the ultraviolet to the microwave. *Appl. Opt.* **23**, 1206–1225 (1984).
14. Mastrapa, R. M. *et al.* Optical constants of amorphous and crystalline H₂O-ice in the near infrared from 1.1 to 2.6 μm . *Icarus* **197**, 307–320 (2008).
15. Mastrapa, R. M. *et al.* Optical constants of amorphous and crystalline H₂O-ice: 2.5–22 μm (4000–455 cm^{-1}) optical constants of H₂O-ice. *Astrophys. J.* **701**, 1347–1356 (2009).
16. Clark, R. N. *et al.* The surface composition of Iapetus: mapping results from Cassini VIMS. *Icarus* **218**, 831–860 (2012).
17. Hapke, B. *Theory of Reflectance and Emittance Spectroscopy* (Cambridge Univ. Press, 2012).
18. Tosi, F. *et al.* Thermal measurements of dark and bright surface features on Vesta as derived from Dawn/VIR. *Icarus* **240**, 36–57 (2014).
19. Fray, N. & Schmitt, B. Sublimation of ices of astrophysical interest: a bibliographic review. *Planet. Space Sci.* **57**, 2053–2080 (2009).
20. Rubin, M. *et al.* Mass transport around comets and its impact on the seasonal differences in water production rates. *Astrophys. J.* **788**, 168 (2014).
21. Crifo, J. F., Loukianov, G. A., Rodionov, A. V. & Zakharov, V. V. Navier-Stokes and direct Monte Carlo simulations of the circumnuclear coma II. Homogeneous, aspherical sources. *Icarus* **163**, 479–503 (2003).
22. De Sanctis, M. C., Lasue, J. & Capria, M. T. Seasonal effects on comet nuclei evolution: activity, internal structure, and dust mantle formation. *Astron. J.* **140**, 1–13 (2010).
23. Rosenberg, E. D. & Prialnik, D. The effect of internal inhomogeneity on the activity of comet nuclei — application to Comet 67P/Churyumov-Gerasimenko. *Icarus* **209**, 753–765 (2010).
24. De Sanctis, M. C. *et al.* Shape and obliquity effects on the thermal evolution of the Rosetta target 67P/Churyumov-Gerasimenko cometary nucleus. *Icarus* **207**, 341–358 (2010).
25. Sunshine, J. M. *et al.* Water ice on comet 103P/Hartley 2. *EPSC-DPS Joint Meeting 2011*, 1345 (2011); available at <http://meetingorganizer.copernicus.org/EPSC-DPS2011/EPSC-DPS2011-1345.pdf>.
26. Vincent, J. B. *et al.* Large heterogeneities in comet 67P as revealed by active pits from sinkhole collapse. *Nature* **523**, 63–66 (2015).

Acknowledgements We thank the following institutions and agencies, which supported this work: the Italian Space Agency (ASI, Italy), Centre National d'Etudes Spatiales (CNES, France), Deutsches Zentrum für Luft- und Raumfahrt (DLR, Germany), and the National Aeronautic and Space Administration (NASA, USA). VIRTIS was built by a consortium from Italy, France and Germany, under the scientific responsibility of the Istituto di Astrofisica e Planetologia Spaziali of the INAF (Italy), which also guides the scientific operations. The VIRTIS instrument development for ESA has been funded and managed by ASI, with contributions from Observatoire de Meudon financed by CNES (France), and from DLR (Germany). We also thank the Rosetta Science Ground Segment and the Rosetta Mission Operations. The VIRTIS calibrated data will be available through the ESA's Planetary Science Archive (PSA) website.

Author Contributions M.C.D.S. and F.C. contributed to data analysis and the writing of the manuscript. G.F. and F.C. provided calibrated VIRTIS data. A.R. and M.C. provided the spectral fit. M.C.D.S., M.T.C., M.F. and S.T. provided the thermal modelling. F.T. retrieved the temperatures. All authors contributed to the discussion of the results and to the writing of the paper.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to M.C.D.S. (mariacristina.desanctis@iaps.inaf.it).

The VIRTIS Team

G. Piccioni¹, G. Bellucci¹, M. Fulchignoni², R. Jaumann³, K. Stephan³, A. Longobardo¹, V. Mennella⁴, A. Migliorini¹, J. Benkhoff⁵, J. P. Bibring⁶, A. Blanco⁷, M. Blecka⁸, R. Carlson⁹, U. Carsenty³, L. Colangeli³, M. Combes², J. Crovisier², P. Drossart², T. Encrenaz², C. Federico¹⁰, U. Fink¹¹, S. Fonti⁷, P. Irwin¹², Y. Langevin⁶, G. Magni¹, L. Moroz³, V. Orofino⁷, U. Schade¹³, F. Taylor¹², D. Tiphene², G. P. Tozzi¹⁴, N. Biver², L. Bonal¹⁵, J.-Ph. Combe¹⁶, D. Despan², E. Flamini¹⁷, S. Fornasier², A. Frigeri¹, D. Grassi¹, M. S. Gudipati⁹, F. Mancarella⁹, K. Markus³, F. Merlin², R. Orosei¹⁸, G. Rinaldi¹, M. Cartacci¹, A. Cicchetti¹, S. Giuppi¹, Y. Hello², F. Henry², S. Jacquino², J. M. Reess², R. Noschese¹, R. Politi¹ & G. Peter³

¹INAF-IAPS, Istituto di Astrofisica e Planetologia Spaziali, 00133 Rome, Italy. ²LESIA, Observatoire de Paris/CNRS/UPMC/Université Paris-Diderot, 92195 Meudon, France. ³Institute for Planetary Research, DLR, 12498 Berlin, Germany. ⁴Osservatorio di Capodimonte, 80131 Napoli, Italy. ⁵European Space Agency – ESTEC, 2201 AZ, The Netherlands. ⁶Institut d'Astrophysique Spatiale CNRS, 91400 Orsay, France. ⁷Dipartimento di Matematica e Fisica "Ennio De Giorgi", Università del Salento, 73100, Lecce, Italy. ⁸Space Research Centre, Polish Academy of Sciences, 00-716 Warsaw, Poland. ⁹NASA JPL, Pasadena, California 91109, USA. ¹⁰Università di Perugia, 06123 Perugia, Italy. ¹¹Lunar Planetary Laboratory, University of Arizona, Tucson, Arizona 85721, USA. ¹²Department of Physics, Oxford University, Oxford OX1 3RH, UK. ¹³Helmholtz-Zentrum Berlin für Materialien und Energie, 14109 Berlin, Germany. ¹⁴Osservatorio Astrofisico di Arcetri, 50125, Firenze, Italy. ¹⁵Université Grenoble Alpes – CNRS Institut de Planetologie et Astrophysique de Grenoble, Batiment D de Physique, BP 53, 38041 Grenoble Cedex 9, France. ¹⁶Bear Fight Institute, Winthrop, Washington 98862, USA. ¹⁷Agenzia Spaziale Italiana, 00133 Rome, Italy. ¹⁸Istituto di Radioastronomia – INAF, 40129 Bologna, Italy.

METHODS

Data. VIRTIS observations were acquired in September 2014 when the Rosetta spacecraft was orbiting at a distance of about 27 km from comet 67P/Churyumov–Gerasimenko’s surface resulting in a spatial resolution on the ground of about 7 m per pixel. During this period the instrument was observing the morning hemisphere with a solar phase of about 60–70°. VIRTIS-M-IR data were acquired in scan mode with an integration time of 3 s. The characteristics of the data here used are reported in Extended Data Table 1.

Spectral modelling. In order to model nucleus surface spectra, a solution of the radiative transfer equation in a particulate medium must be applied. In this paper we adopt the Hapke model¹⁷. Our analysis is performed on normalized spectra in order to rule out the effect of surface roughness and minimize photometric issues. We model the icy regions of the comet, following the equations below, as formed by a two end-members regolith made of a ‘dark terrain’ (DT) and water ice. The DT represents the average spectrum of the comet surface after photometric correction. Water ice is modelled as in ref. 27, starting from optical constants which are obtained from refs 13–16 to cover the VIRTIS wavelength range.

We investigated two mixing modalities: areal and intimate. Areal mixtures are obtained from a linear combination of the reflectances (r) of water ice and DT:

$$r_{\text{eff}} = fr_{\text{H}_2\text{O}} + (1-f)r_{\text{DT}} \quad (1)$$

where f is the relative amount of water ice and r_{eff} is the effective reflectance of the medium. Intimate mixture (‘salt and pepper’) is modelled as a linear combination of the two end-members single scattering albedoes (w) and is given by:

$$w_{\text{eff}} = fw_{\text{H}_2\text{O}} + (1-f)w_{\text{DT}} \quad (2)$$

where the derived w_{eff} is used to compute the final reflectance. In both cases, along with the amount of water, the grain diameter d is also retrieved.

The modelling of the observed spectra is performed by a retrieval procedure that searches for the minimum of the reduced chi-square (χ^2_{R}), namely the best fit between the modelled (r^{m}) and the observed (r^{o}) reflectance:

$$\chi^2_{\text{R}} = \sum_{j=1}^N \left(\frac{r_j^{\text{o}} - r_j^{\text{m}}}{\sigma_j} \right)^2 \frac{1}{\text{DOF}} \quad (3)$$

where j identifies each band (λ), N is the total number of bands, DOF are the degrees of freedom.

The observed spectra are corrected for spikes and instrumental artefacts. Among the residual sources of error, like stray light and signal from the coma, instrumental noise is the main contribution to the error of the measured reflectance (σ_j).

In Extended Data Fig. 1 we report, as an example, the result of a typical spectral fit for a pixel exhibiting a certain amount of ice obtained in the intimate and the areal modes. It can be noted that intimate mixing provides the best match with measured spectra. The reason is that areal mixing increases the relative depth of 1.5- μm and

2- μm absorption bands with respect to the 3- μm feature, while they are very weak or absent across the data set investigated in this work. Given this, the water ice abundance maps are obtained modelling the spectra as intimate mixtures.

The non-detection of the water ice bands at 1.5 μm and 2.0 μm indicates that the water ice and non-icy components are intimately mixed. In fact an areal mixture of 1% of water ice and 99% of non-ice materials also yields spectra with well-defined absorption features at 1.5 and 2 μm as well as an increase in reflectance (Extended Data Fig. 1). This was the case with comet 9P/Tempel 1, where ice-rich patches on the nucleus have been modelled through an areal mixture containing $6 \pm 3\%$ ice².

The maps showing the abundance of water ice (Fig. 3) are produced by setting a threshold of 50 on the median S/N of the spectra in order to avoid pixels in shadow or not of the nucleus of the comet.

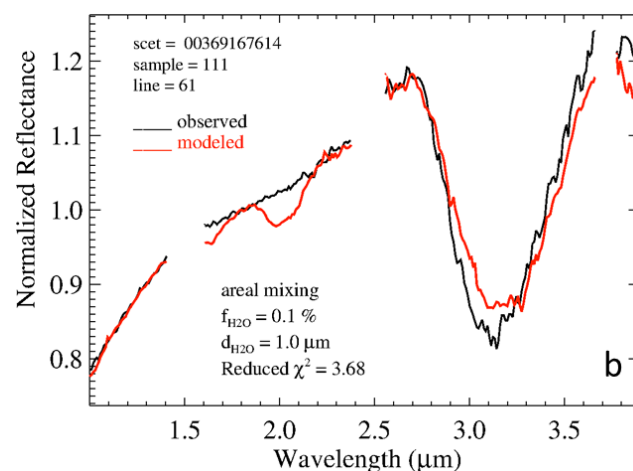
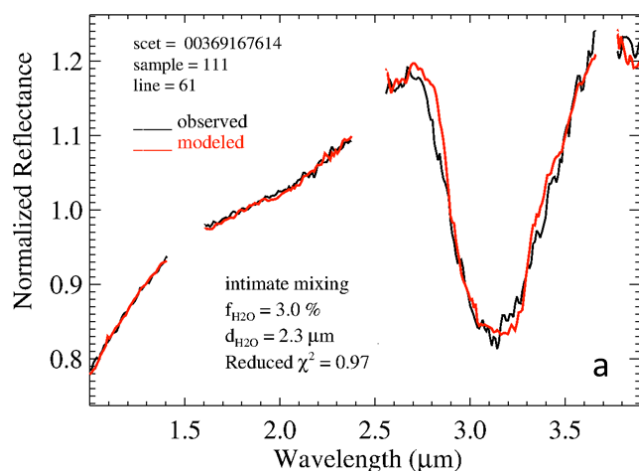
In Extended Data Fig. 2, we report the spectra of three consecutive pixels going out of the shadow, and thus with decreasing abundance of water ice. The relative accuracy of the parameters obtained from the spectral modelling (water ice amount and grain diameter) is of the order of 20%, due to the instrumental noise and the uncertainty on the level of the dark terrain.

Estimation of water flux. The contribution to the total outgassing was estimated using only the extent of ice in our data. We calculated the surface area in the data that contains the transient ice and the percentage of ice in such an area. Using that information, and the pixel size, we computed the equivalent area covered by pure ice. In the VIR data, this area is $\sim 1 \text{ km}^2$.

MIRO measured about 10^{25} molecules s^{-1} and they estimated that between 0.1 and 1% of the 67P nucleus surface is needed to explain the water gas production rates if water ice were located on the surface⁶. Thus, using these values we can say that the ‘transient ice’ seen by VIRTIS contributed $\sim 3\%$ to the total water flux. However, this is the lower limit of the contribution. In fact we can extrapolate that a similar fraction of the neck region, even if not observed by VIR due to observation conditions, is subject to a similar diurnal illumination effect and could have ice deposits like the imaged area.

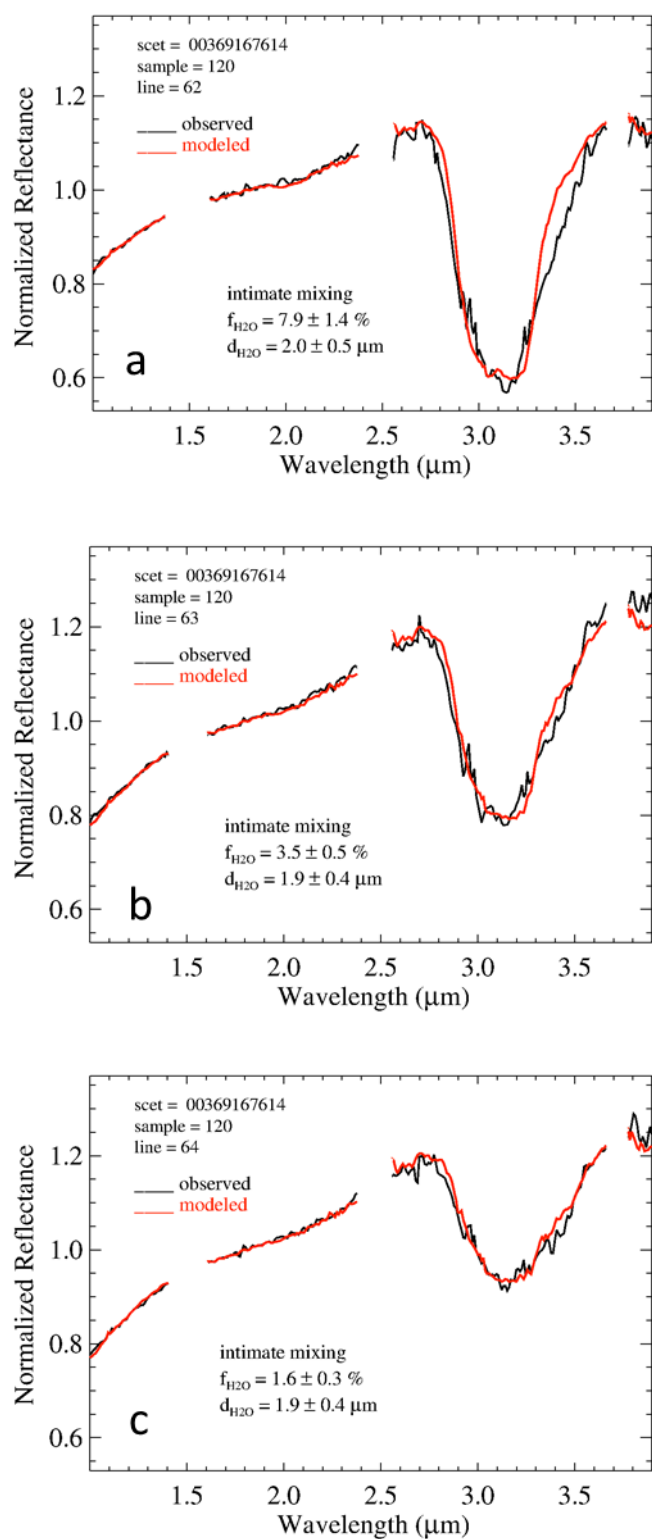
Code availability. The code used to generate the thermal models of comet 67P is a direct implementation of a published model^{24,28,29}. The code used to generate the spectral fit is described above. The code used to retrieve the nucleus temperatures of comet 67P is a direct implementation of a published method¹⁸.

27. Ciarniello, M. *et al.* Hapke modeling of Rhea surface properties through Cassini-VIMS spectra. *Icarus* **214**, 541–555 (2011).
28. De Sanctis, M. C., Capria, M. T. & Coradini, A. Thermal evolution model of 67P/Churyumov-Gerasimenko, the new Rosetta target. *Astron. Astrophys.* **444**, 605–614 (2005).
29. Capria, M. T., Coradini, A., De Sanctis, M. C. & Blecka, M. I. P/Wirtanen thermal evolution: effects due to the presence of an organic component in the refractory material. *Planet. Space Sci.* **49**, 907–918 (2001).
30. JPL Small-Body Database Browser. 67P/Churyumov-Gerasimenko. <http://ssd.jpl.nasa.gov/sbdb.cgi> (NASA/Jet Propulsion Laboratory).

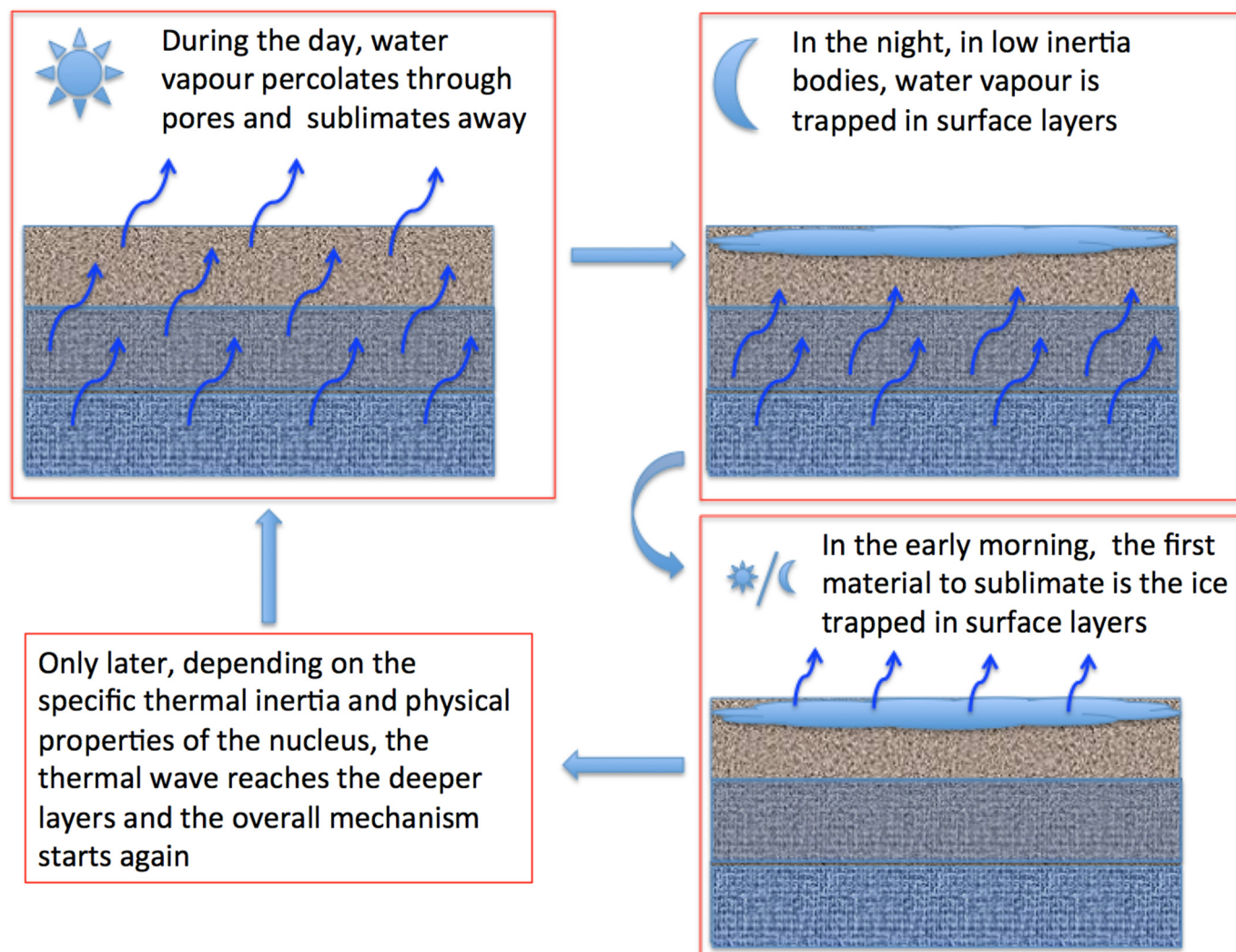


Extended Data Figure 1 | Spectral fit. **a, b**, Two mixing modalities, intimate (**a**) and areal (**b**), are used to model the same spectrum; the spectrum is identified by its position in the acquired image ('sample' and 'line') and its spacecraft event time (scet). The three missing parts of the spectra are related to the wavelength ranges covered by the junctions of the filters which produce significant artefacts. They are thus removed during the fitting procedure. The

spectra are normalized with respect to $\lambda_0 = 1.8 \mu\text{m}$. For the areal mixture case, the modelled absorption band at $2 \mu\text{m}$ is relevant, even with the small abundance ($f_{\text{H}_2\text{O}}$) and grain diameter ($d_{\text{H}_2\text{O}}$) we retrieved. This implies a worse fit, as indicated by the larger χ^2 variable. The intimate mixture is thus a better model of the spectra.

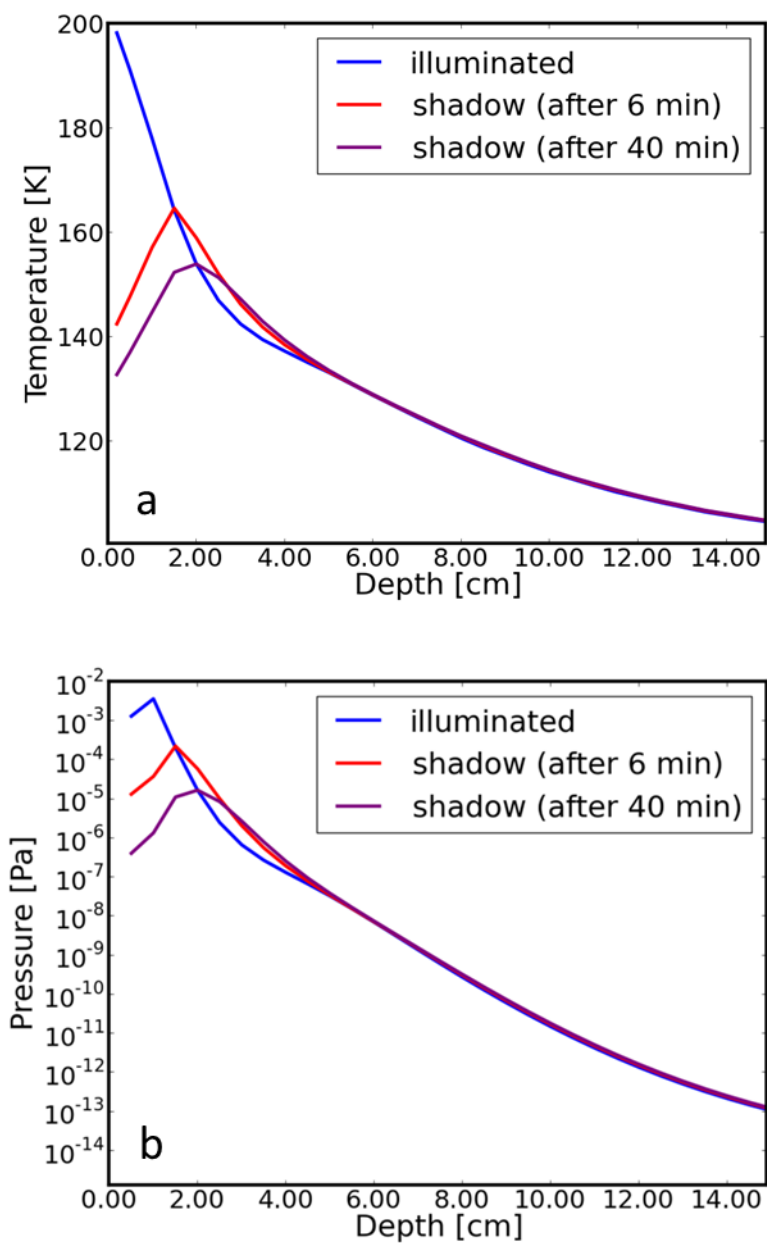


Extended Data Figure 2 | Spectral fits of comet nucleus spectra with different ice content. From a to c, the depth of the absorption band at $3.2 \mu\text{m}$ decreases and the band centre moves slightly towards longer wavelengths. In all cases the spectra are well modelled with a decreasing amount of water ice ($f_{\text{H}_2\text{O}}$) and a constant grain diameter ($d_{\text{H}_2\text{O}}$) of that water ice. Missing parts of the spectra and normalization are as for Extended Data Fig. 1.



Extended Data Figure 3 | Diurnal cycle of water. In the case reported here, the sublimation of water vapour takes place in a deeper layer (see cartoon for a graphic explanation). The water vapour filters up to the surface layers (which are essentially dehydrated, as during the day we do not see any spectral signature of ice) where, finding lower temperature conditions (as in the

night or in the shadow), it condenses and is trapped as ice. The subsequent illumination of the surface leads to absolute loss of the condensed water vapour. This is an effective mechanism of transport of H_2O from deeper layers to the surface.



Extended Data Figure 4 | Temperature and water vapour profiles.

a, Temperature profiles, and **b**, water vapour pressure profiles, from the nucleus surface to the interior at different times: the blue curve is the profile when the

area is illuminated; the red curve is the profile obtained 6 min after passing into shadow; and the purple curve is about 40 min after passing into shadow.

Extended Data Table 1 | Characteristics of VIRTIS observations

FILENAME	II_00369167614.QUB
OBSERVATION START TIME (yyyy-mm-dd hh:mm:ss)	2014-09-12 18:34:43.818
OBSERVATION END TIME (yyyy-mm-dd hh:mm:ss)	2014-09-12 19:10:57.456
SPACECRAFT ALTITUDE	27.849 Km
PHASE ANGLE	63.1 deg
PIXEL RESOLUTION	6.7 m/pix
FILENAME	II_00369210814.QUB
OBSERVATION START TIME (yyyy-mm-dd hh:mm:ss)	2014-09-13 06:34:43.976
OBSERVATION END TIME (yyyy-mm-dd hh:mm:ss)	2014-09-13 07:10:57.604
SPACECRAFT ALTITUDE	28.19 Km
PHASE ANGLE	60.7 deg
PIXEL RESOLUTION	7.1 m/pix
FILENAME	II_00369345274.QUB
OBSERVATION START TIME (yyyy-mm-dd hh:mm:ss)	2014-09-14 19:55:43.981
OBSERVATION END TIME (yyyy-mm-dd hh:mm:ss)	2014-09-14 20:30:57.680
SPACECRAFT ALTITUDE	28.2 Km
PHASE ANGLE	63.6 deg
PIXEL RESOLUTION	7.05 m/pix

Extended Data Table 2 | Main parameters used in the comet model

Physical Quantity	Value	Reference
Rotation Period	12.403 [h]	[12]
Semi-major Axis	3.463 [AU]	[30]
Eccentricity	0.641	[30]
Albedo	0.06	[7]
Emissivity	0.9	[22,24, 28,29]
Dust (Silicates) /Ice	1.5	Adopted for this model and typically used in cometary thermophysical models
Dust (Organics)/Ice	2.0	Adopted for this model and typically used in cometary thermophysical models
Mean Pore Radius	10^{-4} [m]	[22,24]
Porosity	0.6	Adopted for this model and typically used in cometary thermophysical models
Hertz Factor (the area of contact between material grains relative to the cross-sectional area)	0.01	[22,24]
Silicatic Dust Thermal Conductivity	3 [W K ⁻¹ m ⁻¹]	[22,24,29]
Organic Dust Thermal Conductivity	0.25 [W K ⁻¹ m ⁻¹]	[22,24,29]
H ₂ O (crystalline) thermal conductivity	567/T [W K ⁻¹ m ⁻¹]	[22,24,29]
Mean Density	400 [Kg m ⁻³]	Resulting from the parameters used
Thermal Inertia	50-70 [J m ⁻² K ⁻¹ s ^{-0.5}]	Resulting from the parameters used
Initial Temperature	100 [K]	Adopted for this model and typically used in cometary thermophysical models

References are Sierks *et al.*⁷, Mottola *et al.*¹², De Sanctis *et al.*²², De Sanctis *et al.*²⁴, De Sanctis *et al.*²⁸, Capria *et al.*²⁹, JPL Small-Bodies³⁰

References cited in this table are as follows: refs 7, 12, 22, 24, 28, 29, 30.

Controlling neutron orbital angular momentum

Charles W. Clark¹, Roman Barankov², Michael G. Huber³, Muhammad Arif³, David G. Cory^{4,5,6,7} & Dmitry A. Pushin^{6,8}

The quantized orbital angular momentum (OAM) of photons¹ offers an additional degree of freedom and topological protection from noise. Photonic OAM states have therefore been exploited in various applications^{2,3} ranging from studies of quantum entanglement and quantum information science^{4–7} to imaging^{8–12}. The OAM states of electron beams^{13–15} have been shown to be similarly useful, for example in rotating nanoparticles and determining the chirality of crystals^{16–19}. However, although neutrons—as massive, penetrating and neutral particles—are important in materials characterization, quantum information and studies of the foundations of quantum mechanics, OAM control of neutrons has yet to be achieved. Here, we demonstrate OAM control of neutrons using macroscopic spiral phase plates that apply a ‘twist’ to an input neutron beam. The twisted neutron beams are analysed with neutron interferometry. Our techniques, applied to spatially incoherent beams, demonstrate both the addition of quantum angular momenta along the direction of propagation, effected by multiple spiral phase plates, and the conservation of topological charge with respect to uniform phase fluctuations. Neutron-based studies of quantum information science^{20,21}, the foundations of quantum mechanics^{22,23}, and scattering and imaging²⁴ of magnetic, superconducting and chiral materials have until now been limited to three degrees of freedom: spin, path and energy. The optimization of OAM control, leading to well defined values of OAM, would provide an additional quantized degree of freedom for such studies.

OAM is associated with rotation of a particle about a fixed axis. The axial particle currents are encoded in the spiral phase profile of the particle’s wavefunction. The component of OAM parallel to the axis of rotation is quantized in integer multiples of the reduced Planck constant \hbar . The quantization of OAM is a consequence of the wavefunction having a single value, which implies that it is a periodic function of the rotation angle, with a period of 2π radians. When interactions of the particle with its environment are symmetric with respect to such rotations, OAM is conserved. Changes in OAM can be effected by putting a ‘twist’ on the wavefunction. To achieve this with neutrons, we use a macroscopic phase plate made in the shape of a spiral staircase, which matches the phase profile of an OAM state. Neutrons passing through such a spiral phase plate (SPP) obtain axial rotation around the direction of the beam (the quantization axis), dictated by a specific design of the plate, as we discuss below. Similar techniques have been applied previously to light², X-rays²⁵ and electrons¹³.

The production and detection of neutron OAM states is difficult for two reasons. First, even for the cold neutrons used in our experiments, it is almost impossible to get two-dimensional (2D) neutron imaging detectors with resolution better than the micrometre-scale neutron coherence length. Second, our neutron source is a nuclear reactor with thermal and cryogenic moderators, spectrally filtered with Bragg-scattering monochromators. This makes it difficult to produce a spatially coherent neutron beam and directly detect an eigenstate of OAM. We estimate that the transverse coherence length of the neutron beam in our experiments is in the range from 60 nm (ref. 26) to a few

micrometres (ref. 23) depending on the transverse direction within the beam profile, while the beam diameter is about 15 mm.

The input to the interferometer (Fig. 1a) contains a mixture of OAM states and is spatially incoherent over the transverse displacement of the neutron paths. We adopt the strategy of using neutron interferometry to demonstrate coherent control of self-interfering neutrons. At any given time there is at most one neutron in the interferometer. How an SPP changes the angular-momentum composition of a wavefunction Ψ , can be thought of as follows. Upon transmission through the SPP, Ψ is simply modulated by the transmission amplitude $\Psi \rightarrow \exp(i\theta)\Psi$, where θ is the phase function of the SPP. Figure 1b is a schematic diagram of an SPP, and Fig. 1c is a photograph of two such plates (see details in Methods). As shown in Fig. 1b, the thickness of a SPP, h , varies with the angle, φ , so that $h = h_0 + h_s\varphi/(2\pi)$, where h_0 is the base height of the SPP and h_s is the step height. Interaction of low-energy neutrons with materials can be described using an optical potential²³. Within this formalism, the phase shift with respect to vacuum of a neutron passing through an SPP is:

$$\theta = -Nb_c\lambda h = -Nb_c\lambda\left(h_0 + h_s\frac{\varphi}{2\pi}\right) \quad (1)$$

where N is the atom density and b_c is the coherent scattering length of the material composing the phase plate, and $\lambda = 0.271$ nm is the

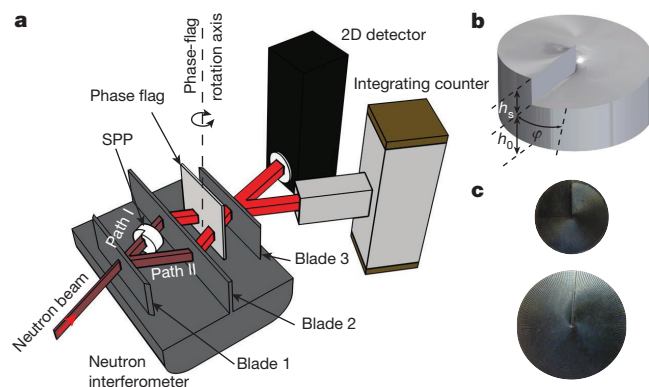


Figure 1 | Schematic diagram of the neutron interferometer. **a**, The input neutron beam is coherently split into two coherent paths by the Bragg diffraction at blade 1. Blade 2 serves as a neutron mirror and blade 3 recombines the neutron paths and directs them to two neutron detectors: an integrating counter and a two-dimensional imaging detector. The neutron counts recorded by the two detectors contain information about the relative phase of the neutron wavefunction accumulated along the two separate paths. The phase flag is a 2-mm-thick fused silica plate. By positioning and rotating the phase flag we can adjust a uniform phase difference between the neutron paths inside the interferometer. The SPP is placed in one path to produce a spatial phase distribution across the neutron wavefront. **b**, Schematic diagram of the SPP. The step height of the spiral, h_s , is chosen to match a 2π phase shift difference between a path passing the plate and the reference path in air. **c**, Photographs of actual SPPs. Two SPPs that produce a 4π phase shift ($h_s = 224$ μm), photographed with the spiral staircase rising clockwise on the front face: one is 10 mm in diameter (top), the other is 15 mm in diameter (bottom).

¹Joint Quantum Institute, National Institute of Standards and Technology and University of Maryland, Gaithersburg, Maryland 20899, USA. ²Photonics Center and Department of Biomedical Engineering, Boston University, Massachusetts 02215, USA. ³National Institute of Standards and Technology, Gaithersburg, Maryland 20899, USA. ⁴Department of Chemistry, University of Waterloo, Waterloo, Ontario N2L 3G1, Canada. ⁵Perimeter Institute for Theoretical Physics, Waterloo, Ontario N2L 2Y5, Canada. ⁶Institute for Quantum Computing, University of Waterloo, Waterloo, Ontario N2L 3G1, Canada. ⁷Canadian Institute for Advanced Research, Toronto, Ontario M5G 1Z8, Canada. ⁸Department of Physics and Astronomy, University of Waterloo, Waterloo, Ontario N2L 3G1, Canada.

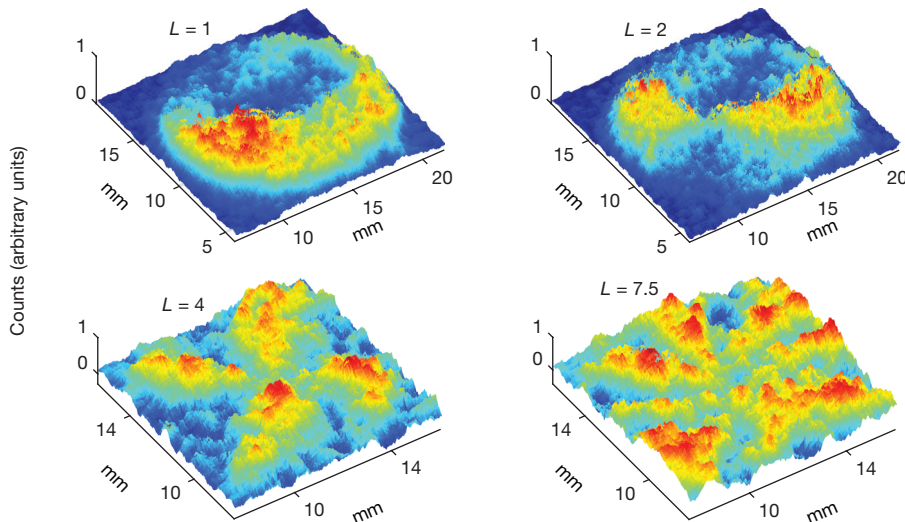


Figure 2 | OAM interferograms. Spatial distribution of the neutron counts in the 2D detector³⁰ of the neutron interferometer for four SPPs, with values of $L = 1, 2, 4$ and 7.5 , as labelled. The horizontal and vertical positions on the 2D neutron detector are shown in millimetres. For the integer values of L these distributions display the simple OAM interference pattern expressed in equation (2); for $L = 7.5$ we have the superposition of OAM modes given by equation (4) in Methods. The 2D detector is a centroid-type event-counting detector with a spatial resolution of $100\ \mu\text{m}$ and an 18% quantum efficiency (that is, counts registered per neutron incident on the detector). Its operation is shot-noise (Poisson-noise) limited in this regime. The neutron counts were collected over 3.5 days and normalized by the maximal pixel count, which is about 45.

neutron wavelength. To attain a 2π phase shift in an aluminium SPP, the step height is $h_s = 112\ \mu\text{m}$. Note that this value of h_s is much greater than our neutron wavelength, since the index of refraction, n , for neutrons at this wavelength in aluminium is very close to unity²⁷: $1 - n = Nb_c\lambda^2/(2\pi) \approx 2.43 \times 10^{-6}$. Figure 1c shows the photographs of two SPPs with $h_s = 224\ \mu\text{m}$.

Immediately after passing the SPP a neutron acquires an additional azimuthal phase twist of $2\pi L$, where the effective angular momentum $L = Nb_c\lambda h_s/(2\pi)$ is a function of the spiral step height. Every OAM component of an incoherent beam entering the phase plate obtains an additional phase twisting proportional to L . The average angular momentum of the neutron, measured in units of \hbar , becomes $\langle L \rangle = L_0 + L$, where L_0 is the initial average OAM, which varies for different neutrons of the incoherent beam. Given the uncertainty of the angular momentum with which a neutron enters the phase plate, the outgoing states also remain uncertain.

The neutron OAM states that we generate are analysed using interferometry. We fabricated several SPPs, corresponding to phase circulation of $2\pi, 4\pi, 8\pi$ and 15π around the singularity of the wavefunction, or average orbital momenta of $L = 1, 2, 4$ and 7.5 . The input neutron wavefunction is split along two paths in the neutron interferometer. Both neutron paths pass through a 2-mm-thick plate (the phase flag), with which we control the relative overall phase difference between paths. One neutron path passes through an SPP, which imprints an azimuthal phase upon the neutron wavefront. Up to a common overall phase, the neutron wavefunction at the entrance of the 2D detector can be described in cylindrical coordinates centred at the axis of the phase plate as: $\Psi = (c_1 e^{iL\varphi} + c_2 e^{-i\phi_0})\Psi_0$, where φ is the azimuthal angle about the beam propagation axis (see Fig. 1b), c_1 and c_2 are amplitudes composed of neutron reflection and transmission coefficients of the interferometer blades, ϕ_0 is the phase due to the phase flag and Ψ_0 is the wavefunction of a neutron entering the interferometer. The spatially resolved neutron fluence rate at the 2D detector is thus proportional to:

$$I_{2D}(\rho, \varphi; \theta) \propto [1 + C \cos(L\varphi + \phi_0)] |\Psi_0|^2 \quad (2)$$

where $0 \leq C \leq 1$ is the interferometric contrast of the interferometer. For the interferometer used in this work, $C = 0.84$ without background correction.

Figure 1a is a schematic diagram of the experiment performed at the Neutron Interferometry and Optics Facility (NIOF)^{28,29} at the National Institute of Standards and Technology (NIST). The neutron interferometer shown there has the standard configuration used in the field²³: it is analogous to an optical Mach–Zehnder interferometer. Created in a 20-MW research reactor, cooled in a cryogenic moderator to a temperature of 20 K, and transported through 30 m of neutron guides, the neutrons enter the neutron interferometer in a

15-mm-diameter beam, which is Bragg-scattered on the first blade of the neutron interferometer into two coherent paths. We insert one or two SPPs into one of the paths, so a neutron in that path acquires a variation of phase across its wavefront. The second blade of the neutron interferometer is employed as a lossy mirror: part of the incident beam is transmitted through the blade and leaves the neutron interferometer (this part is not shown in Fig. 1a); the remainder of the beam is Bragg-scattered towards the third blade of the neutron interferometer. The two paths from the second blade reconnect coherently at the third blade. This blade combines the interfering transmitted/Bragg-diffracted neutron paths and directs them into the 2D detector³⁰ and an integrating ^3He neutron counter²⁸. A 2-mm-thick plate of fused silica is installed between the second and third blades. This serves as a ‘phase flag’: by rotating this phase flag we can introduce and control a spatially uniform phase difference between neutron paths. The spatially resolved data from the 2D detector provides information on the spatial phase ‘impressed’ upon on the neutron wavefront by the SPP.

Figure 2 shows 2D images using a position-sensitive detector placed after the third blade of the neutron interferometer (see Fig. 1a). These interferograms use a conventional false-colour spatial representation of the time-integrated neutron intensity per pixel of the 2D detector (examples of the raw data, data manipulation and data noise are shown in Extended Data Figs 1–3). The results generated by SPPs with $h_s = 112\ \mu\text{m}$, $224\ \mu\text{m}$, $448\ \mu\text{m}$ and $840\ \mu\text{m}$, which correspond to $L = 1$,

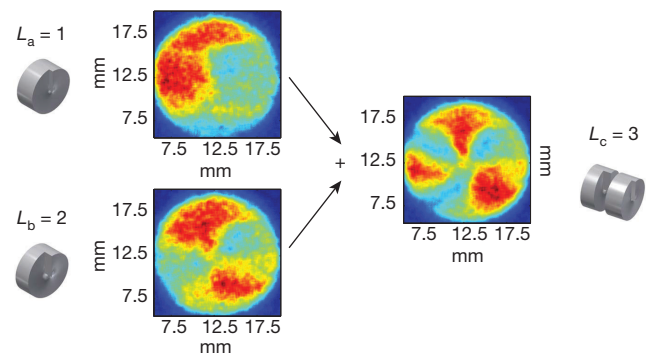


Figure 3 | Addition of angular momenta along the direction of propagation accomplished by two spiral phase plates. On the left are two interferograms obtained separately with different SPPs: with effective OAM $L_a = 1$ ($h_s = 112\ \mu\text{m}$) and with $L_b = 2$ ($h_s = 224\ \mu\text{m}$). The horizontal and vertical positions on the 2D neutron detector are shown in millimetres. On the right is the interferogram obtained by concatenating both SPPs in path I of the neutron interferometer (in the configuration indicated schematically), to make a compound SPP with $L_c = L_a + L_b = 3$.

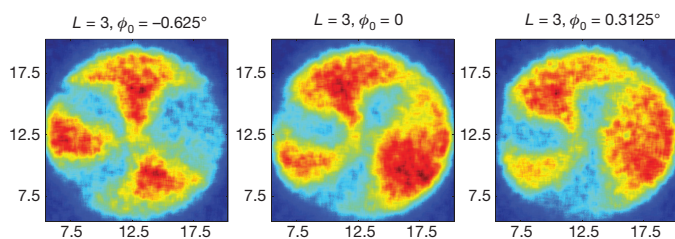


Figure 4 | Rotational invariance. Interferograms obtained using the compound SPP with $L = 3$, as described in Fig. 3 for different positions of the phase flag (see Fig. 1a). The horizontal and vertical positions on the 2D neutron detector are shown (in millimetres). The three images were taken at different settings of the phase flag, corresponding to the values $\phi_0 = -0.625^\circ$, 0° and 0.3215° , as indicated. The image rotates by the angle proportional to the phase flag. The average OAM is independent of the phase flag position, that is, L is preserved.

2, 4 and 7.5, are shown in Fig. 2. If the effective thickness L is close to an integer value, it follows from equations (4) and (5) in Methods that the interferogram is dominated by the characteristic pattern corresponding to the OAM $l \approx L$. Thus for integer values of L we see the expected sharp angular distribution of neutron intensity, dominated by the contribution from $l \approx L$, while for non-integer L we have a distributed superposition of contributions from numerous values of l . We find that the action of multiple SPPs follows the familiar rule for addition of angular momenta. If we concatenate two SPPs with step heights corresponding to different OAMs L_a and L_b , as shown in Fig. 3, the net transfer of OAM to the neutron beam is $L_c = L_a + L_b$. Figure 3 thus provides an experimental demonstration of the elementary proposition $1 + 2 = 3$. Moreover, the composite angular momentum L_c transforms as a true angular momentum under rotations of the composite phase plate structure about its symmetry axis, as well as with respect to interference with the beam in the other arm of the neutron interferometer (Fig. 4).

In conclusion, we have demonstrated control of OAM of neutrons using easily manufactured macroscopic SPPs. The average OAM of the beams has been measured using a perfect-crystal neutron interferometer. The interferometric experiments exemplify the celebrated particle–wave duality of neutrons: on the one hand, neutrons are detected as individual particles, while on the other, neutrons traverse space like waves, carrying quantized values of OAM. Our interferograms indicate that each of these individual states has its OAM changed by the same amount when passed through a SPP of integer order. We showed that the action of two phase plates placed in series increased the OAM by the sum of the OAM produced by each phase plate separately. This result is consistent with the accepted notion of additivity of OAM, and is also consistent with the conservation of vortex topological charge. Possible future directions of such studies include optimizing SPPs and experimenting with alternative phase-shifting devices, such as the fork-dislocation gratings also used in OAM control of electrons and photons; increasing the coherence of the input neutron beam; entanglement of neutron spin and orbital angular momenta; and angular-momentum-resolved measurements of neutron scattering. OAM states may also extend phase-contrast neutron imaging²⁸ to two dimensions. Even in its nascent form, as demonstrated here, neutron OAM control in a single crystal interferometer provides an entangling operation between path and OAM. This potentially offers a new tool for neutron-based quantum information science and tests of the foundations of quantum mechanics^{22,23}.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 20 April; accepted 26 July 2015.

- Allen, L., Beijersbergen, M. W., Spreeuw, R. J. C. & Woerdman, J. P. Orbital angular momentum of light and the transformation of Laguerre-Gaussian laser modes. *Phys. Rev. A* **45**, 8185–8189 (1992).

- Yao, A. M. & Padgett, M. J. Orbital angular momentum: origins, behavior and applications. *Adv. Opt. Photon.* **3**, 161–204 (2011).
- Torres, J. P. & Torner, L. *Twisted Photons: Applications of Light with Orbital Angular Momentum* (Wiley-VCH, 2011).
- Mair, A., Vaziri, A., Weihs, G. & Zeilinger, A. Entanglement of the orbital angular momentum states of photons. *Nature* **412**, 313–316 (2001).
- Vallone, G. *et al.* Free-space quantum key distribution by rotation-invariant twisted photons. *Phys. Rev. Lett.* **113**, 060503 (2014).
- Yu, S. Potentials and challenges of using orbital angular momentum communications in optical interconnects. *Opt. Express* **23**, 3075 (2015).
- Cardano, F. *et al.* Quantum walks and wavepacket dynamics on a lattice with twisted photons. *Sci. Adv.* **1**, e1500087 (2015).
- Hell, S. W. Far-field optical nanoscopy. *Science* **316**, 1153–1158 (2007).
- Bahrndt, J. *et al.* First observation of photons carrying orbital angular momentum in undulator radiation. *Phys. Rev. Lett.* **111**, 034801 (2013).
- Takahashi, Y. *et al.* Bragg x-ray ptychography of a silicon crystal: visualization of the dislocation strain field and the production of a vortex beam. *Phys. Rev. B* **87**, 121201 (2013).
- Gariépy, G. *et al.* Creating high-harmonic beams with controlled orbital angular momentum. *Phys. Rev. Lett.* **113**, 153901 (2014).
- Fischer, P. X-ray imaging of magnetic structures. *IEEE Trans. Magn.* **51**, 0800131 (2014).
- Uchida, M. & Tonomura, A. Generation of electron beams carrying orbital angular momentum. *Nature* **464**, 737–739 (2010).
- Verbeeck, J., Tian, H. & Schattschneider, P. Production and application of electron vortex beams. *Nature* **467**, 301–304 (2010).
- McMorran, B. J. *et al.* Electron vortex beams with high quanta of orbital angular momentum. *Science* **331**, 192–195 (2011).
- Schattschneider, P., Löffler, S., Stöger-Pollach, M. & Verbeeck, J. Is magnetic chiral dichroism feasible with electron vortices? *Ultramicroscopy* **136**, 81–85 (2014).
- Rusz, J., Bhowmick, S., Eriksson, M. & Karlsson, N. Scattering of electron vortex beams on a magnetic crystal: towards atomic-resolution magnetic measurements. *Phys. Rev. B* **89**, 134428 (2014).
- Juchtmans, R., Béché, A., Abakumov, A., Batuk, M. & Verbeeck, J. Using electron vortex beams to determine chirality of crystals in transmission electron microscopy. *Phys. Rev. B* **91**, 094112 (2015).
- Shiloh, R. *et al.* Unveiling the orbital angular momentum and acceleration of electron beams. *Phys. Rev. Lett.* **114**, 096102 (2015).
- Pushin, D. A., Huber, M. G., Arif, M. & Cory, D. G. Experimental realization of decoherence-free subspace in neutron interferometry. *Phys. Rev. Lett.* **107**, 150401 (2011).
- Wood, C. J. *et al.* Quantum correlations in a noisy neutron interferometer. *Phys. Rev. A* **90**, 032315 (2014).
- Klepp, J., Sponar, S. & Hasegawa, Y. Fundamental phenomena of quantum mechanics explored with neutron interferometers. *Prog. Theor. Exp. Phys.* **2014**, 082A01 (2014).
- Rauch, H. & Werner, S. A. *Neutron Interferometry: Lessons in Experimental Quantum Mechanics, Wave-Particle Duality, and Entanglement* 2nd edn, Vol. 12 (Oxford Univ. Press, 2015).
- Pfeiffer, F. in *Neutron Imaging Applications: A Reference for the Imaging Community* (eds Anderson, I. S. *et al.*) Ch. 8 (Springer, 2009).
- Peele, A. G. *et al.* X-ray phase vortices: theory and experiment. *J. Opt. Soc. Am. A* **21**, 1575–1584 (2004).
- Pushin, D. A., Arif, M., Huber, M. G. & Cory, D. G. Measurements of the vertical coherence length in neutron interferometry. *Phys. Rev. Lett.* **100**, 250404 (2008).
- Granada, J. Slow neutron total cross-section of Al6061 at low temperatures. *J. Nucl. Mater.* **277**, 346–350 (2000).
- Pushin, D. A., Cory, D. G., Arif, M., Jacobson, D. L. & Huber, M. G. Reciprocal space approaches to neutron imaging. *Appl. Phys. Lett.* **90**, 224104 (2007).
- Pushin, D. A. *et al.* Neutron interferometry at the National Institute of Standards and Technology. *Adv. High Energy Phys.* **2015**, 687480 (2015).
- Dietze, M., Felber, J., Raum, K. & Rausch, C. Intensified CCDs as position sensitive neutron detectors. *Nucl. Instrum. Meth. A* **377**, 320–324 (1996).

Acknowledgements Financial support provided by the NSERC ‘Create’ and ‘Discovery’ programmes, CERC and the NIST Quantum Information Program is acknowledged. We appreciate discussions with B. McMorran, D. Sarenac, S. Werner and K. Wright.

Author Contributions C.W.C. and D.A.P. conceived the idea of implementing OAM for neutron beams. D.A.P. conceived the idea of using the spiral phase plate and neutron interferometer for OAM implementation, and designed and built the experimental setup. M.A., M.G.H. and D.A.P. conducted the neutron interferometer experiments. R.B., C.W.C., D.G.C. and D.A.P. led the analysis and wrote the manuscript, with M.A. and M.G.H. contributing substantially.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to D.A.P. (dmitry.pushin@uwaterloo.ca).

METHODS

Experimental methods. Our experiments were performed at the NIST Center for Neutron Research in Gaithersburg, USA. Neutrons are produced by a 20-MW split-core reactor moderated with heavy water, and are further cooled to around $T = 20$ K by a liquid hydrogen cold source. The NIOF facility is about 25 m from the reactor core, and neutrons are guided to it by a multilayer neutron guide. At NIOF, neutrons are extracted from the guide by a pyrolytic graphite monochromator and collimated by a set of cadmium slits. The neutron wavelength $\lambda = 0.271$ nm (energy $E \approx 11$ meV) used in the experiment was chosen to satisfy the Bragg condition ($\theta_B = 25.6^\circ$) for the (111) crystallographic planes of a crystal silicon interferometer. Figure 1a shows the neutron interferometer used in this experiment, which has the Mach–Zehnder configuration that is in standard use in this field²³.

The integrating ^3He detector was used for aligning the system, monitoring of the reactor flux, and determining the experimental parameters such as the initial interferometer phase (ϕ_0) and interferometer contrast (C). This phase and contrast can be measured by rotating the phase flag (see angle of rotation in Fig. 1a) inside the interferometer with no sample present. During experiments with samples we use the integrating detector to monitor reactor neutron intensity.

The 2D imaging neutron detector³⁰ has a spatial resolution of 100 μm and detection efficiency of 18% at the neutron wavelength of $\lambda = 0.271$ nm. The count rate over the whole area of the 2D detector was approximately 1.9 neutrons per second and a typical measurement time for each image collection was 3.5 days. These long collection times necessitate robust phase stability of the neutron interferometer platform. The interferometer phase drift was 1° per day and the image noise per pixel was statistically limited. Images shown in Figs 2–4 were obtained with the 2D detector, filtered with standard ‘average’ filter over 10×10 pixels, and normalized to a maximal count. The raw typical images of the detector without filtering and normalization can be seen in Extended Data Figs 1–3. The integrating counter in our experiment is a ^3He detector with nearly 100% efficiency.

The SPPs were machined out of dowels of Al 6061 alloy by a 5-axis computer numerical control milling machine at the NIST machine shop, following two different procedures. The smaller plates (10 mm diameter) were cut by rotating the Al sample and moving the end mill out, while the larger plates (15 mm in diameter) were milled in the form of a helical staircase with approximately 200 treads.

The experimental setup (that is, neutron interferometer with samples and phase flag, and the neutron detectors) is located within three nested enclosures (Matryoshka-style). To minimize phase drifts during the week-long data collection times, the temperature of the innermost enclosure is actively controlled to remain at 24°C within 5 mK (ref. 26). The middle enclosure is a Faraday cage with temperature isolation and sound damping. It sits on a 40,000-kg, vibration-isolated table suspended on air springs from a platform decoupled from the floor of the reactor facility³¹. The vibration isolation actively suppresses the mechanical noise spectrum above 0.5 Hz and is controlled with micrometre precision. Changing the SPPs requires opening of all three nested enclosures, and we found it necessary to wait about 24 h afterwards for the system to return to equilibrium.

Theoretical analysis. OAM states of neutrons can be described with the paraxial approximation of wave mechanics. Within this approximation, the wavefunction of a freely propagating neutron can be written as $\Phi(\mathbf{r}) = e^{ik_z z} \Psi(\mathbf{r})$, where $\mathbf{r} = (x, y, z)$ is the position vector, and the envelope function $\Psi(\mathbf{r})$ satisfies the 2D Helmholtz equation:

$$\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \right) \Psi + 2ik_z \frac{\partial \Psi}{\partial z} = 0 \quad (3)$$

where $k_z \approx \sqrt{2mE}/\hbar$ is the wavevector of a neutron with mass m and energy E propagating forward along the z axis. OAM states comprise a complete set of solutions of this equation¹ in the cylindrical coordinates (ρ, ϕ, z) , where $x = \rho \cos(\phi)$, and $y = \rho \sin(\phi)$. The eigenfunctions take the form $\Psi(\rho, \phi, z) = u_l(\rho, z) \exp(i l \phi)$, where the function $u_l(\rho, z)$ describes the transverse radial structure of the wavefront as a function of the propagation coordinate z . The function $\exp(i l \phi)$ is an eigenfunction of the OAM operator, $\hat{L}_z = -i\hbar \partial / \partial \phi$, with eigenvalue $l\hbar$, where $l = 0, \pm 1, \pm 2, \dots$ is an integer. In the simplest case when $l = 0$, the envelope function reduces to a diffracting beam having a Gaussian profile in the transverse direction at any axial position z .

For an arbitrary value of L the phase factor $e^{iL\phi}$ generated by an SPP describes a superposition of OAM states:

$$e^{iL\phi} = \sum_{l=0, \pm 1, \dots} \beta_l e^{il\phi} \quad (4)$$

where the amplitudes β_l are given by the overlaps of the phase factor and the OAM eigenfunctions:

$$\beta_l = e^{i(L-l)\pi} \text{sinc}(L-l) \quad (5)$$

where $\text{sinc}(x) = \sin(\pi x)/(\pi x)$. For an OAM beam with zero angular momentum passing through the SPP L , the probability distribution of the resulting OAM states is provided by $w_l = |\beta_l|^2$. We notice that for a non-integer effective thickness of the phase plate, such as $L = 7.5$, as used in our experiment, a superposition of different OAM states would be produced with dominant contributions from $l = 7$ and $l = 8$. In general, given the uncertainty of L , even for the values of $L = n + \delta L$ close to the integers $n = 0, \pm 1, \dots$, the resulting state is a mixture of OAM states.

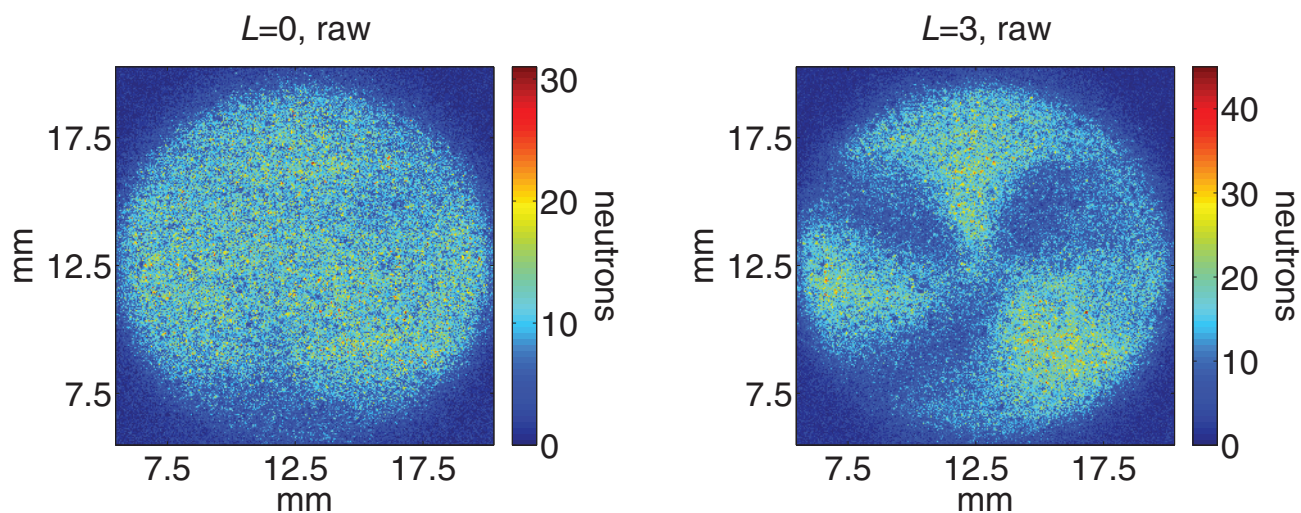
The distribution of the OAM states in any beam experiment depends upon the coherence properties of the incident beam. This issue has recently been discussed in the context of electron vortex experiments by Shiloh *et al.*¹⁹. In our experiment, the coherence length in the direction orthogonal to the propagation axis is $l_c \approx 60$ nm (the corresponding wavevector component $k_c = 1/l_c \approx 17 \mu\text{m}^{-1}$). Since the beam radius $R = 7.5$ mm is much larger than the coherence length, the beam is spatially incoherent, and the neutron beam as a whole does not have a definite state of angular momentum.

Future applications and experiments.

Studies of quantum information science and foundational questions of quantum mechanics. OAM states of neutrons could be used to extend the size of the Hilbert space available in neutron studies of fundamental quantum phenomena. Neutrons have been useful testbeds for quantum information with coherent degrees of freedom demonstrated for path, spin and momentum^{32–38}. Many of the most useful demonstrations and applications of quantum information are richer in higher dimensions, including quantum error correction²⁰ and contextuality.

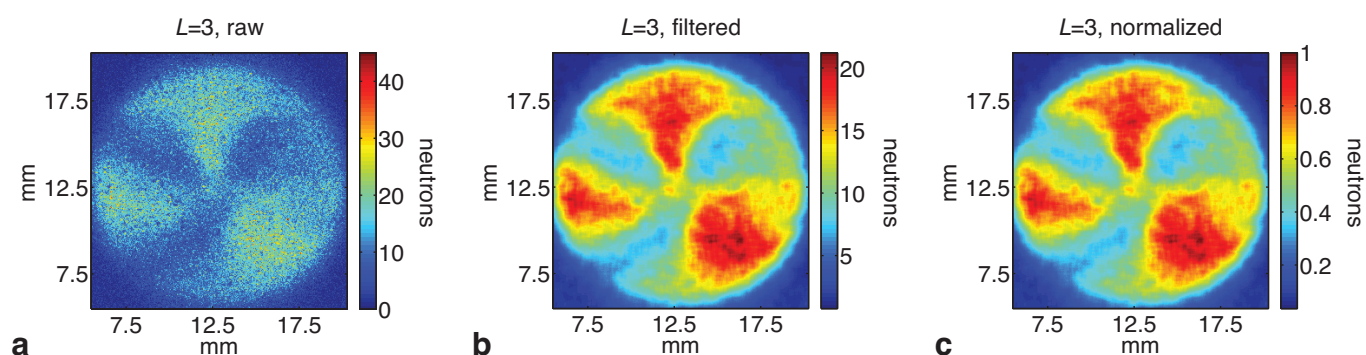
Neutron OAM for materials characterization. Recently demonstrated neutron imaging and tomography^{39,40} paved a way to many applications²⁴, such as studies of magnetic materials^{41,42}, magnetic domains in bulk materials^{43,44} and also superconductors⁴⁵. OAM states of neutrons may benefit those studies. Indeed, the OAM state may provide a unique probe for samples having embedded chirality such as chiral nanoparticles and chiral liquid crystals. The neutron microscope^{46,47} that is being currently developed at the NIST Center for Neutron Research has a potential to use neutron OAM beams (similar to ref. 48) to study magnetic materials and chiral structures. A combination of OAM and spin states of neutrons may also be employed to study skyrmions, which are currently investigated using small-angle scattering techniques⁴⁹ or magnetic reflectometry.

- Arif, M., Brown, D. E., Greene, G. L., Clothier, R. & Littrell, K. Multistage position-stabilized vibration isolation system for neutron interferometry. *Proc. SPIE Int. Soc. Opt. Eng.* **2264**, 20–26 (1994).
- Hasegawa, Y., Loidl, R., Badurek, G., Baron, M. & Rauch, H. Violation of a Bell-like inequality in single-neutron interferometry. *Nature* **425**, 45–48 (2003).
- Bartosik, H. *et al.* Experimental test of quantum contextuality in neutron interferometry. *Phys. Rev. Lett.* **103**, 040403 (2009).
- Sponar, S., Klepp, J., Zeiner, C., Badurek, G. & Hasegawa, Y. Violation of a Bell-like inequality for spin-energy entanglement in neutron polarimetry. *Phys. Lett. A* **374**, 431–434 (2010).
- Hasegawa, Y. *et al.* Entanglement between degrees of freedom of single neutrons. *Nucl. Instrum. Methods Phys. A* **611**, 310–313 (2009).
- Hasegawa, Y. *et al.* Engineering of triply entangled states in a single-neutron system. *Phys. Rev. A* **81**, 032121 (2010).
- Sponar, S. *et al.* High-efficiency manipulations of triply entangled states in neutron polarimetry. *New J. Phys.* **14**, 53032 (2012).
- Sponar, S. *et al.* Triple entanglement in neutron interferometric and polarimetric experiments. *J. Phys. G* **340**, 012044 (2012).
- Pfeiffer, F. *et al.* Neutron phase imaging and tomography. *Phys. Rev. Lett.* **96**, 215505 (2006).
- Strobl, M. *et al.* Neutron dark-field tomography. *Phys. Rev. Lett.* **101**, 123902 (2008).
- Kardjilov, N. *et al.* Three-dimensional imaging of magnetic fields with polarized neutrons. *Nature Phys.* **4**, 399–403 (2008).
- Grünzweig, C. *et al.* Bulk magnetic domain structures visualized by neutron dark-field imaging. *Appl. Phys. Lett.* **93**, 112504 (2008).
- Manke, I. *et al.* Three-dimensional imaging of magnetic domains. *Nature Commun.* **1**, 125 (2010).
- Kardjilov, N., Manke, I., Hilger, A., Strobl, M. & Banhart, J. Neutron imaging in materials science. *Mater. Today* **14**, 248–256 (2011).
- Treimer, W. Radiography and tomography with polarized neutrons. *J. Magn. Magn. Mater.* **350**, 188–198 (2014).
- Liu, D. *et al.* Demonstration of achromatic cold-neutron microscope utilizing axisymmetric focusing mirrors. *Appl. Phys. Lett.* **102**, 183508 (2013).
- Liu, D. *et al.* Demonstration of a novel focusing small-angle neutron scattering instrument equipped with axisymmetric mirrors. *Nature Commun.* **4**, 2556 (2013).
- Verbeeck, J. *et al.* Atomic scale electron vortices for nanoresearch. *Appl. Phys. Lett.* **99**, 203109 (2011).
- Mühlbauer, S. *et al.* Skyrmion lattice in a chiral magnet. *Science* **323**, 915–919 (2009).

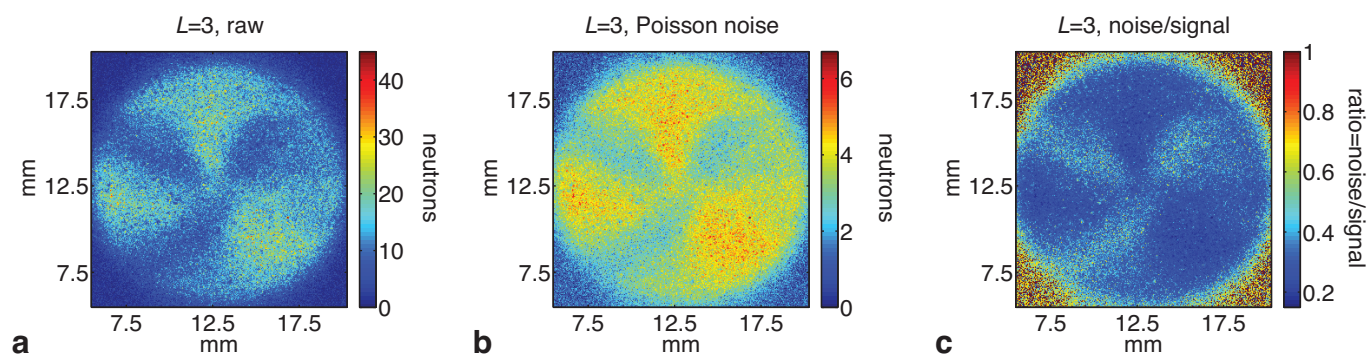


Extended Data Figure 1 | Raw data. Typical images of raw neutron count data obtained over about 80 contiguous hours of data collection with the 2D imaging detector (see Fig. 1). False-colour representation of neutron counts per pixel as indicated by scale on image. On the left is an image of an SPP with

$L = 0$ that is equivalent to a uniform phase plate; on the right is an image of the $L = 3$ compound SPP discussed in Fig. 3. The horizontal and vertical positions on the 2D neutron detector are shown in millimetres.



Extended Data Figure 2 | Data processing. Illustrations of steps taken to convert raw images collected on the 2D detector to the images shown in Figs 2–4. **a**, Raw data. **b**, Same data, passed through 2D averaging filter with averaging taken over a 10 pixel \times 10 pixel square. **c**, Filtered data normalized to maximum value of intensity in **b**. The horizontal and vertical positions on the 2D neutron detector are shown in millimetres.



Extended Data Figure 3 | Noise distribution. Illustrations of steps taken to model effects of shot noise in the raw images collected on the 2D detector and images shown in Figs 2–4. **a**, Image of the raw data. **b**, Poisson noise or

square root of each pixel shown in **a**. **c**, Noise-to-signal ratio of image in **a**. The horizontal and vertical positions on the 2D neutron detector are shown in millimetres.

A concise synthesis of (+)-batzelladine B from simple pyrrole-based starting materials

Brendan T. Parr¹, Christos Economou¹ & Seth B. Herzon¹

Alkaloids, secondary metabolites that contain basic nitrogen atoms, are some of the most well-known biologically active natural products in chemistry and medicine¹. Although efficient laboratory synthesis of alkaloids would enable the study and optimization of their biological properties², their preparation is often complicated by the basicity and nucleophilicity of nitrogen, its susceptibility to oxidation, and its ability to alter reaction outcomes in unexpected ways—for example, through stereochemical instability and neighbouring group participation. Efforts to address these issues have led to the invention of a large number of protecting groups that temper the reactivity of nitrogen³; however, the use of protecting groups typically introduces additional steps and obstacles into the synthetic route. Alternatively, the use of aromatic nitrogen heterocycles as synthetic precursors can attenuate the reactivity of nitrogen and streamline synthetic strategies⁴. Here we use such an approach to achieve a synthesis of the complex anti-HIV alkaloid (+)-batzelladine B in nine steps (longest linear sequence) from simple pyrrole-based starting materials. The route uses several key transformations that would be challenging or impossible to implement using saturated nitrogen heterocycles and highlights some of the advantages of beginning with aromatic reagents.

The retrosynthetic conversion of a saturated nitrogen heterocycle to a heteroaromatic exchanges a reactive, basic functional group with one that is lower in energy, non-basic, non-nucleophilic, and more easily manipulated. For example, analysis of well-appreciated physical organic scales of basicity and nucleophilicity shows that the six-membered heterocycle piperidine is much more basic ($pK_b = 3.1$, DMSO, ref. 5) and nucleophilic (nucleophilicity parameter $N = 18.1$, H_2O , ref. 6) than is the corresponding aromatic heterocycle pyridine ($pK_b = 10.6$, DMSO, ref. 7; $N = 11.0$, H_2O , ref. 6). Moreover, functionalized heteroaromatics are readily elaborated by well-established carbon–carbon bond-forming reactions, such as cross-couplings. In the strategy we pursue here, simple pyrrole-based precursors serve as sources of partially or fully saturated nitrogen heterocycles and are advanced by carbon–carbon bond-forming and reductive transformations. This approach complements terpene synthesis and biosynthesis, which typically proceeds by oxidation of a complex hydrocarbon template⁸.

We applied this strategy towards a synthesis of the guanidinium alkaloid (+)-batzelladine B (**1**, Fig. 1a)⁹. Structurally, **1** contains a *syn*-tricyclic guanidine (vessel) connected to a bicyclic guanidine (anchor) via an alkyl ester. At least 15 batzelladine alkaloids have been isolated^{9–12} and several members of this family inhibit the binding of HIV glycoprotein gp120 to human CD4 receptor cells (half-maximal inhibitory concentration of **1** is $31 \mu M$)⁹, thereby preventing viral induction. The absolute stereochemistries of the vessel and anchor of **1** were established in refs 13 and 14, respectively, and syntheses and synthetic studies of other batzelladines have been reported (for selected examples, see refs 15–22). Notably, in ref. 13, a tethered Biginelli condensation strategy was developed and this has provided access to several batzelladines and related alkaloids²³. In refs 16 and 17,

enantioselective synthetic routes to (+)-batzelladine A (**2**) are reported, but a route to **1** has not been described.

We envisioned that the vessel and anchor fragments of **1** (Fig. 1b) could be derived from the pyrrole-based precursors **3** and **6**, respectively, if suitable methods for carbon–carbon bond formation and controlled reduction in oxidation state could be achieved. A rhodium-catalysed formal [4 + 3] cycloaddition between **3** and a donor–acceptor carbene²⁴ was proposed to provide entry to the dehydrotropane **4**, which contains all of the functional group handles required for synthesis of the vessel fragment. We posited that the pyrrole **6** could serve as a precursor to the anchor of **1** by a Mannich addition to form **7** (ref. 25), followed by cyclization and controlled adjustment of oxidation state, with concomitant isomerization.

The *N*-amidinylpyrrole **3** (Fig. 2a) was prepared in two steps and 75% yield from commercial reagents (see Supplementary Information). Extensive experimentation was required to realize the formal [4 + 3] cycloaddition with high yield and stereoselectivity. Ultimately, we found that use of the (*S*)-pantolactonyl α -diazo ester **9** (ref. 26) and dirhodium(II) tetrakis[*N*-phthaloyl-(*S*)-*tert*-leucinate] ($Rh_2[(S)\text{-pttl}]_4$) as catalyst (0.5 mol%) provided the dehydrotropane **10** in 93% yield and >95:5 diastereoselectivity. Formal cycloaddition between **3** and

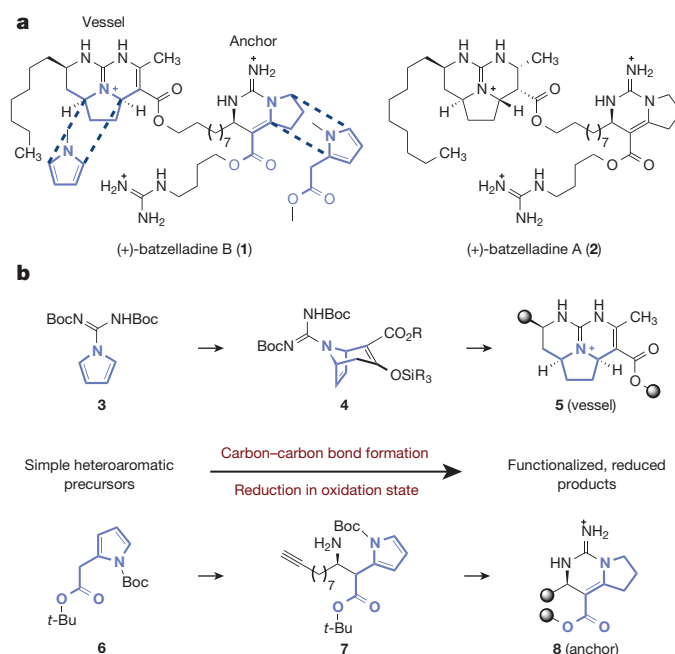


Figure 1 | Structure and synthetic analysis of (+)-batzelladine B (**1**). **a**, The chemical structures of (+)-batzelladine B (**1**) and (+)-batzelladine A (**2**), with embedded pyrrole substructures shown. **b**, The strategy we used produces the vessel and anchor substructures of **1** (**5** and **8**, respectively) from the pyrrole-based starting materials **3** and **6**, via the intermediates **4** and **7**.

¹Department of Chemistry, Yale University, New Haven, Connecticut 06520, USA.

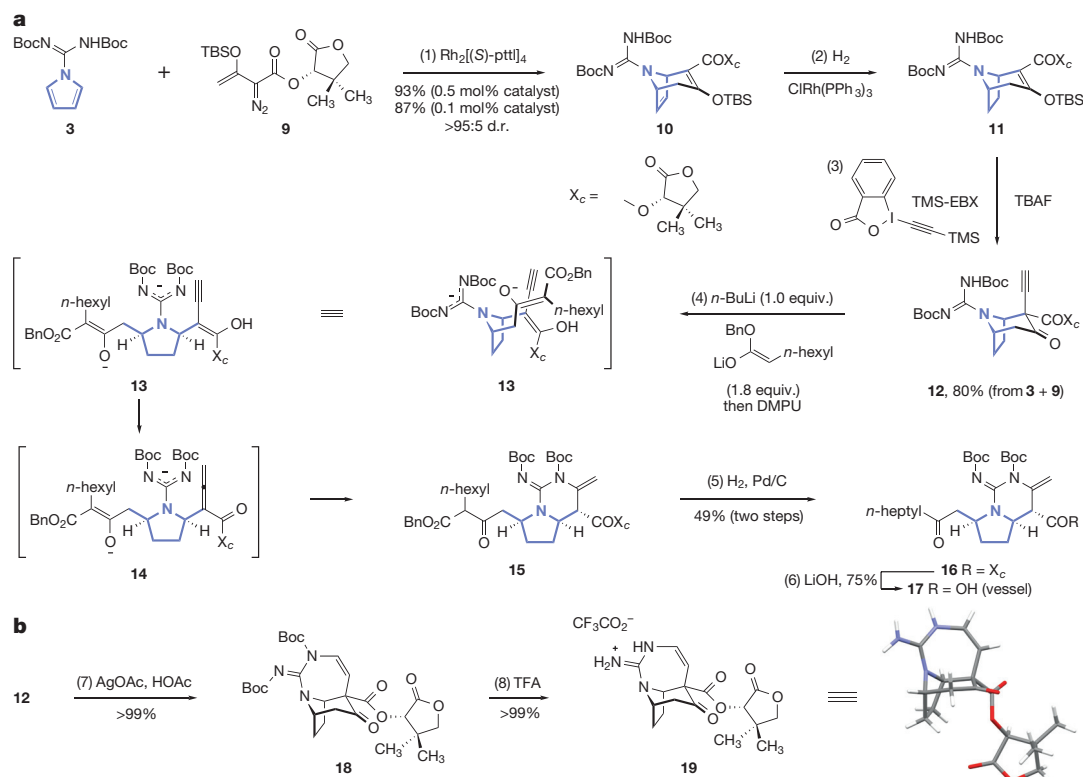


Figure 2 | Synthesis of the vessel fragment of (+)-batzelladine B (1) and determination of its stereochemistry. a, Synthesis of the vessel precursor 17. Reagents and conditions: (1) $\text{Rh}_2[(\text{S})\text{-pttl}]_4$ (0.5 mol%), pentane, 36 °C, 93%, >95:5 d.r., or $\text{Rh}_2[(\text{S})\text{-pttl}]_4$ (0.1 mol%), pentane, 36 °C, 87%, >95:5 d.r.; (2) H_2 (30 atm), $\text{ClRh}(\text{PPh}_3)_3$ (2.0 mol%), *i*-PrOH, 23 °C; (3) tetra-*n*-butylammonium fluoride (TBAF), TMS-EBX, THF- CH_2Cl_2 (8:1), -78 °C,

80% (from 3 + 9); (4) *n*-BuLi, then lithium benzyl octanoate, THF, then DMPU, -78 °C; (5) H_2 (1 atm), Pd/C (10 mol%), THF, 23 °C, 49% (two steps); and (6) LiOH, THF- H_2O (2:1), 0 °C, 75%. **b,** The relative stereochemistry of 12 was established by cyclization and deprotection, followed by X-ray analysis. Reagents and conditions: (7) AgOAc, AcOH, CH_2Cl_2 , 24 °C, >99% and (8) TFA, CH_2Cl_2 , 0–23 °C, >99%.

ent-9 using the same catalyst provided 10 with 76:24 diastereoselectivity (81% yield), demonstrating that the former substrate–catalyst pair is stereochemically matched (an example of double asymmetric synthesis²⁷). Formal cycloaddition between 3 and achiral diazoesters afforded the corresponding adducts in 45%–93% yield and 60%–86% enantiomeric excess (Supplementary Table 1). The yield of 10 was essentially unaffected (87%) when the catalyst loading was reduced to 0.1 mol%. With this key step accomplished, the pyrroline ring was selectively reduced by treatment with chlorotris(triphenylphosphine)rhodium under dihydrogen (10 → 11). Exposure of the reduction product 11 to *n*-tetrabutylammonium fluoride and 1-[(trimethylsilyl)ethynyl]-1,2-benziodoxol-3(1*H*)-one (TMS-EBX)²⁸ at -78 °C formed the α -alkynyl- β -ketoester 12 as a single diastereomer (¹H NMR analysis). The first three steps of this sequence were readily telescoped to provide 12 in 80% overall yield after one purification. The relative stereochemistry of 12 was unequivocally established by 7-*endo*-dig hydroguanylation²⁹ (12 → 18, Fig. 2b) followed by carbamate cleavage (18 → 19) and X-ray analysis.

We then investigated the ring-opening of the bicyclic skeleton of 12 by cleavage of the β -ketoester. Because 12 presents four acidic sites, a careful balance between the protonation state of the substrate and the basicity of the incoming nucleophile was essential to achieving the desired mode of reactivity. After intensive experimentation and optimization, we found that deprotonation of 12 with *n*-butyllithium (1.0 equiv.) followed by the addition of lithium benzyl octanoate (1.8 equiv.) afforded the bicyclic pyrrolidine 15. This cascade sequence is thought to proceed by 1,2-addition to the β -ketoester, retro-aldol ring-opening, and proton transfer to provide the enolone 13. Isomerization of 13 to the acylallene 14 followed by Michael addition of the guanidinyll anion and neutralization of the resulting enolate may then provide 15. The addition of 1,3-dimethyl-3,4,5,6-tetrahydro-2-

pyrimidinone (DMPU) was necessary to promote the retro-aldol ring-opening. Other nucleophiles, such as sodium methoxide, morpholine, ethanethiol, and organometallic reagents (for example, Grignard, organocerium, organotitanium, or organozinc reagents), were also investigated, but in most instances complex mixtures of products were obtained. The addition–rearrangement product 15 exists as a mixture of diastereomers and tautomers; consequently, the β -ketoester of the unpurified product was cleaved with palladium on carbon under dihydrogen, to form the ketone 16 (49% from 12). Saponification of the pantolactonyl ester (lithium hydroxide) afforded the keto-acid 17 (75%; 29% overall from 3).

The anchor fragment was assembled by the sequence shown in Fig. 3, beginning with a highly-diastereoselective Mannich addition²⁵ to form the β -aminoester of the target. Treatment of 6 with lithium diisopropylamide (LDA) and chloro tris(isopropoxy)titanium, followed by addition of the sulfinimine 20, provided the product 21 in 99% yield. The addition product 21 was formed as a single detectable C2 stereoisomer and an inconsequential (approximately 94:6) mixture of C1 stereoisomers (¹H NMR analysis). The C1 and C2 stereocentres were assigned by analogy to related products²⁵ and the C2 stereochemistry was confirmed by derivatization (see Supplementary Information). Notably, attempts to functionalize saturated analogues of 6 by a Mannich addition would be complicated by issues of diastereoselectivity and β -elimination. Owing to the presence of the alkyne and the difficulties associated with handling the vinylogous carbamate of the target¹⁷, reduction of the pyrrole ring was postponed until later in the sequence. The *tert*-butanesulfinyl substituent of 21 was cleaved by treatment with hydrochloric acid in methanol, and the resulting product was cyclized in the presence of bis(chlorodibutyltin)oxide to provide the urea 22 (78%, two steps). *O*-Selective ethylation formed an *iso*-urea (90%, not shown) that was treated with 2,4-(dimethoxy)benzyl

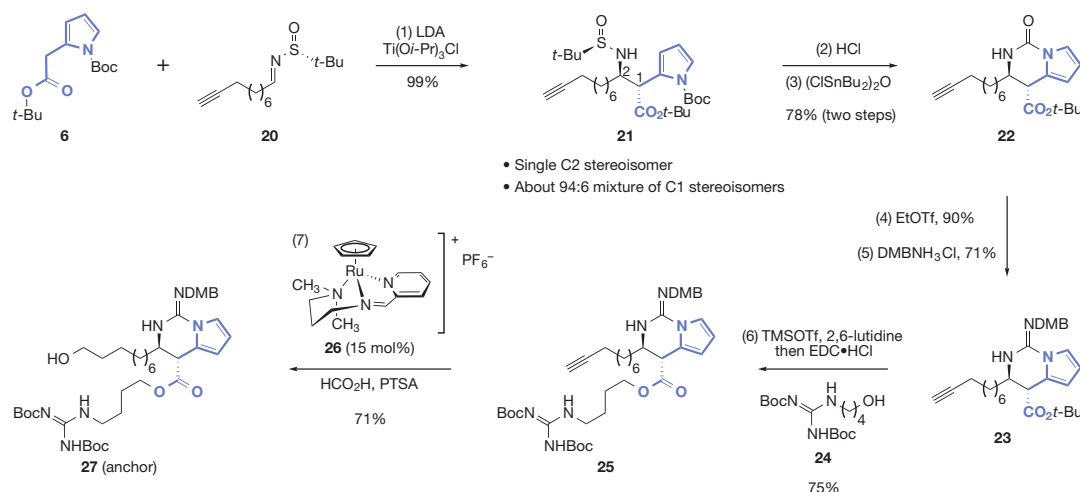


Figure 3 | Synthesis of the anchor fragment of (+)-batzelladine B

(1). Reagents and conditions: (1) LDA, $\text{Ti}(\text{O}i\text{-Pr})_3\text{Cl}$, THF, -78°C , 99%, >20:1 mixture of C1 stereoisomers, approximately 94:6 mixture of C2 stereoisomers; (2) HCl, CH_3OH –1,4-dioxane (4.4:1), 0°C ; (3) $(\text{ClSnBu}_2)_2\text{O}$, toluene, 100°C , 78% (two steps); (4) EtOTf, 2,4,6-tri-*tert*-butyl-pyrimidine,

CH_2Cl_2 , 23°C , 90%; (5) DMBNH_3Cl , 3 Å molecular sieves, EtOH, 70°C , 71%; (6) TMSOTf, 2,6-lutidine, CH_2Cl_2 , 0 – 23°C , then **24**, 4-(dimethylamino)-pyridine (DMAP), 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide (EDC)·HCl, CH_2Cl_2 , 0 – 23°C , 75%; and (7) **26** (15 mol%), PTSA (1.0 equiv.), HCO_2H , *N*-methyl-2-pyrrolidinone (NMP)– H_2O (4:1), 23°C , 71%.

(DMB) amine hydrogen chloride to afford the guanidine **23** (71%). The ester was then cleaved (trimethylsilyl trifluoromethanesulfonate, 2,6-lutidine) and the resultant carboxylic acid was coupled with the alcohol **24** to provide **25**, which contains the complete carbon framework of the anchor (75%). Anti-Markovnikov reductive hydration³⁰ of the terminal alkyne of **25** mediated by the ruthenium catalyst **26** (15 mol%) formed the alcohol **27** (71%; 26% overall from **6**). The addition of *p*-toluenesulfonic acid (PTSA) to quantitatively protonate the guanidine was essential in this step; in its absence, the conversion of **25** was low.

The vessel and anchor fragments **17** and **27** were coupled using 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide hydrochloride (EDC·HCl) to provide the penultimate intermediate **28** (77%), and the synthesis was completed by the carefully optimized sequence shown in Fig. 4. A dry mixture of palladium on carbon and the coupling product **28** was suspended in trifluoroacetic acid under argon for 2 h at 24°C . Under these conditions, the four *tert*-butoxycarbonyl protecting groups were cleaved, the liberated vessel domain underwent cyclodehydration,

and the 1,1-disubstituted enamide was isomerized into conjugation with the ester (**28** → **29**). Upon completion of this step (as judged by ultra-high-performance liquid chromatography/mass spectrometry analysis), the atmosphere within the reaction vessel was replaced with dihydrogen. Stirring the resultant mixture for 18 h at 24°C effected stereoselective reduction of the trisubstituted eneguanidine of the vessel (>20:1 diastereometric ratio (d.r.); see Supplementary Information)¹⁵, controlled semireduction of the anchor pyrrole with tandem isomerization of the resultant dihydropyrrole, and cleavage of the DMB substituent, to provide **1** in 40% isolated yield (45% by NMR).

Previous approaches to synthesizing batzelladine alkaloids and related natural products have used non-aromatic (aliphatic) nitrogen precursors, followed by stepwise adjustments (typically, increases) of oxidation state. The approach we have presented proceeds in the opposite direction and begins with oxidized nitrogen heteroaromatics, followed by carbon–carbon bond-forming reactions and controlled reduction to achieve the saturation patterns of the target. This

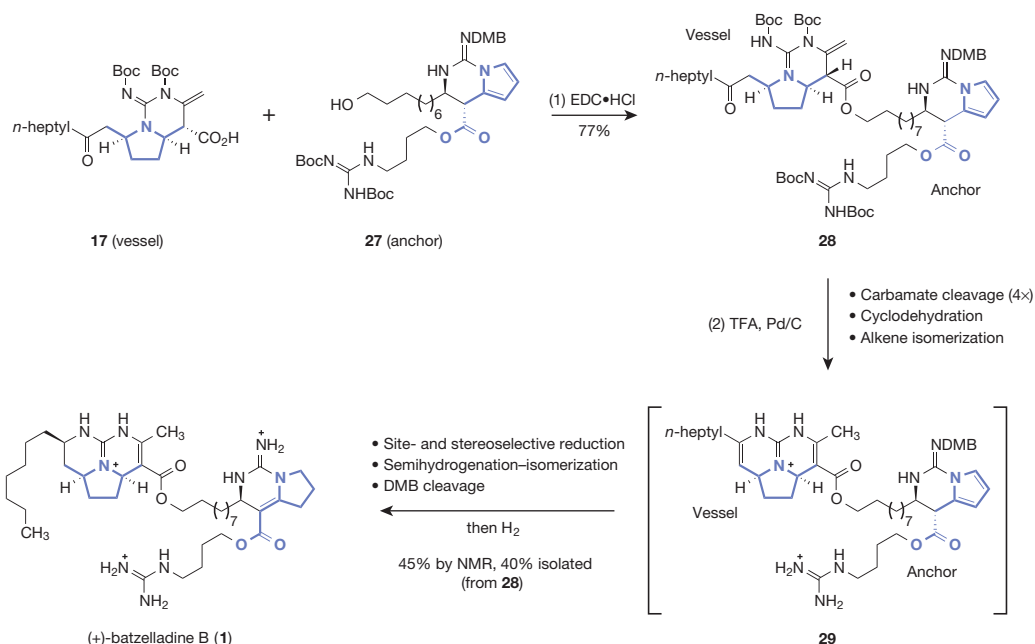


Figure 4 | Coupling of **17 and **27** and completion of the synthesis of (+)-batzelladine B (**1**).** Reagents and conditions: (1) EDC·HCl, DMAP, CH_2Cl_2 , 24°C , 77% and (2) TFA, Pd/C, argon, 0°C , then H_2 , 24°C , 45% (NMR), 40% (isolated).

approach demonstrates additional synthetic pathways that are not apparent or viable when starting from aliphatic nitrogen building blocks, and tempers nitrogen's promiscuous and often problematic reactivity. An added virtue of this strategy is its dependence on late-stage carbon–hydrogen bond-forming reactions, which are among the most reliable classes of transformations.

Received 13 May; accepted 10 July 2015.

Published online 16 September 2015.

- Roberts, M.F. & Wink, M. (eds) *Alkaloids: Biochemistry, Ecology, and Medicinal Applications* (Plenum, 1998).
- Wender, P. A. & Miller, B. L. Synthesis at the molecular frontier. *Nature* **460**, 197–201 (2009).
- Greene, T. W. & Wuts, P. G. M. *Protective Groups in Organic Synthesis* 3rd edn, Ch. 7 (Wiley, 1999).
- Lipshutz, B. H. Five-membered heteroaromatic rings as intermediates in organic synthesis. *Chem. Rev.* **86**, 795–819 (1986).
- Crampton, R. M. & Robotham, I. A. Acidities of some substituted ammonium ions in dimethyl sulfoxide. *J. Chem. Res. Synop.* 22–23 (1997).
- Mayr, H. & Ofial, A. R. Kinetics of electrophile–nucleophile combinations: a general approach to polar organic reactivity. *Pure Appl. Chem.* **77**, 1807–1821 (2005).
- Bordwell, F. G. Equilibrium acidities in dimethyl sulfoxide solution. *Acc. Chem. Res.* **21**, 456–463 (1988).
- Chen, K. & Baran, P. S. Total synthesis of eudesmane terpenes by site-selective C–H oxidations. *Nature* **459**, 824–828 (2009).
- Patil, A. D. et al. Novel alkaloids from the sponge *Batzella* sp.: inhibitors of HIV gp120–human CD4 binding. *J. Org. Chem.* **60**, 1182–1188 (1995).
- Patil, A. D. et al. Batzelladines F–I, novel alkaloids from the sponge *Batzella* sp.: inducers of p56^{lck}–CD4 dissociation. *J. Org. Chem.* **62**, 1814–1819 (1997).
- Hua, H.-M. et al. Batzelladine alkaloids from the Caribbean sponge *Monanchora unguifera* and the significant activities against HIV-1 and AIDS opportunistic infectious pathogens. *Tetrahedron* **63**, 11179–11188 (2007).
- Laville, R. et al. Bioactive guanidine alkaloids from two Caribbean marine sponges. *J. Nat. Prod.* **72**, 1589–1594 (2009).
- Franklin, A. S., Ly, S. K., Mackin, G. H., Overman, L. E. & Shaka, A. J. Application of the tethered Biginelli reaction for enantioselective synthesis of batzelladine alkaloids. Absolute configuration of the tricyclic guanidine portion of batzelladine B. *J. Org. Chem.* **64**, 1512–1519 (1999).
- Duron, S. G. & Gin, D. Y. Synthesis and determination of absolute configuration of the bicyclic guanidine core of batzelladine A. *Org. Lett.* **3**, 1551–1554 (2001).
- Snider, B. B., Chen, J., Patil, A. D. & Freyer, A. J. Synthesis of the tricyclic portions of batzelladines A, B and D. Revision of the stereochemistry of batzelladines A and D. *Tetrahedron Lett.* **37**, 6977–6980 (1996).
- Shimokawa, J., Shirai, K., Tanatani, A., Hashimoto, Y. & Nagasawa, K. Enantioselective total synthesis of batzelladine A. *Angew. Chem. Int. Ed.* **43**, 1559–1562 (2004).
- Arnold, M. A., Day, K. A., Durón, S. G. & Gin, D. Y. Total synthesis of (+)-batzelladine A and (–)-batzelladine D via [4 + 2]-annulation of vinyl carbodiimides with *N*-alkyl imines. *J. Am. Chem. Soc.* **128**, 13255–13260 (2006).
- Cohen, F. & Overman, L. E. Evolution of a strategy for the synthesis of structurally complex batzelladine alkaloids. Enantioselective total synthesis of the proposed structure of batzelladine F and structural revision. *J. Am. Chem. Soc.* **128**, 2594–2603 (2006).
- Cohen, F. & Overman, L. E. Enantioselective total synthesis of batzelladine F and definition of its structure. *J. Am. Chem. Soc.* **128**, 2604–2608 (2006).
- Evans, P. A., Qin, J., Robinson, J. E. & Bazin, B. Enantioselective total synthesis of the polycyclic guanidine-containing marine alkaloid (–)-batzelladine D. *Angew. Chem. Int. Ed.* **46**, 7417–7419 (2007).
- Butters, M. et al. Synthesis and stereochemical determination of batzelladine C methyl ester. *Org. Biomol. Chem.* **7**, 5001–5009 (2009).
- Babji, N. R. & Wolfe, J. P. Asymmetric total synthesis of (+)-merobatzelladine B. *Angew. Chem. Int. Ed.* **51**, 4128–4130 (2012).
- Aron, Z. D. & Overman, L. E. The tethered Biginelli condensation in natural product synthesis. *Chem. Commun.* 253–265 (2004).
- Reddy, R. P. & Davies, H. M. L. Asymmetric synthesis of tropanes by rhodium-catalyzed [4 + 3] cycloaddition. *J. Am. Chem. Soc.* **129**, 10312–10313 (2007).
- Tang, T. P. & Ellman, J. A. Asymmetric synthesis of β -amino acid derivatives incorporating a broad range of substitution patterns by enolate additions to *tert*-butanesulfinyl imines. *J. Org. Chem.* **67**, 7819–7832 (2002).
- Davies, H. M. L., Ahmed, G. & Churchill, M. R. Asymmetric synthesis of highly functionalized 8-oxabicyclo[3.2.1]octene derivatives. *J. Am. Chem. Soc.* **118**, 10774–10782 (1996).
- Masamune, S., Choy, W., Petersen, J. S. & Sita, L. R. Double asymmetric synthesis and a new strategy for stereochemical control in organic synthesis. *Angew. Chem. Int. Ed.* **24**, 1–30 (1985).
- Fernández González, D., Brand, J. P. & Waser, J. Ethynyl-1,2-benziodoxol-3(1*H*)-one (EBX): an exceptional reagent for the ethynylation of keto, cyano, and nitro esters. *Chem. Eur. J.* **16**, 9457–9461 (2010).
- Gainer, M. J., Bennett, N. R., Takahashi, Y. & Looper, R. E. Regioselective rhodium(II)-catalyzed hydroaminations of propargylguanidines. *Angew. Chem. Int. Ed.* **50**, 684–687 (2011).
- Zeng, M., Li, L. & Herzon, S. B. A highly active and air-stable ruthenium complex for the ambient temperature anti-Markovnikov reductive hydration of terminal alkynes. *J. Am. Chem. Soc.* **136**, 7058–7067 (2014).

Supplementary Information is available in the online version of the paper.

Acknowledgements Financial support from the National Institutes of Health (NRSA fellowship GM110898-01A1 to B.T.P., Chemistry Biology Interface Training Program T32GM067543 to C.E.) and Yale University is acknowledged. We thank B. Mercado for X-ray crystallographic analysis of **19** and K.-p. Wang for assistance with HPLC purification of **1**.

Author Contributions B.T.P. and C.E. performed and analysed the experiments. All authors contributed to the design of experiments and composition of the manuscript.

Author Information Crystallographic data for **19** have been deposited at the Cambridge Crystallographic Data Centre as CCDC 1400311. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to S.B.H. (seth.herzon@yale.edu).

Lithospheric controls on magma composition along Earth's longest continental hotspot track

D. R. Davies¹, N. Rawlinson², G. Iaffaldano^{1†} & I. H. Campbell¹

Hotspots are anomalous regions of volcanism at Earth's surface that show no obvious association with tectonic plate boundaries. Classic examples include the Hawaiian–Emperor chain and the Yellowstone–Snake River Plain province. The majority are believed to form as Earth's tectonic plates move over long-lived mantle plumes: buoyant upwellings that bring hot material from Earth's deep mantle to its surface¹. It has long been recognized that lithospheric thickness limits the rise height of plumes^{2–4} and, thereby, their minimum melting pressure. It should, therefore, have a controlling influence on the geochemistry of plume-related magmas, although unambiguous evidence of this has, so far, been lacking. Here we integrate observational constraints from surface geology, geochronology, plate-motion reconstructions, geochemistry and seismology to ascertain plume melting depths beneath Earth's longest continental hotspot track, a 2,000-kilometre-long track in eastern Australia that displays a record of volcanic activity between 33 and 9 million years ago^{5,6}, which we call the Cosgrove track. Our analyses highlight a strong correlation between lithospheric thickness and magma composition along this track, with: (1) standard basaltic compositions in regions where lithospheric thickness is less than 110 kilometres; (2) volcanic gaps in regions where lithospheric thickness exceeds 150 kilometres; and (3) low-volume, leucite-bearing volcanism in regions of intermediate lithospheric thickness. Trace-element concentrations from samples along this track support the notion that these compositional variations result from different degrees of partial melting, which is controlled by the thickness of overlying lithosphere. Our results place the first observational constraints on the sub-continental melting depth of mantle plumes and provide direct evidence that lithospheric thickness has a dominant influence on the volume and chemical composition of plume-derived magmas.

Plate tectonic theory successfully describes how the lithosphere—Earth's rigid outermost shell—consists of a mosaic of segments that move and interact across the surface of our planet. It also accounts for the majority of Earth's volcanism, which is concentrated at plate boundaries. However, an important class of volcanism occurs within plates or across plate boundaries, often forming linear volcanic chains that grow older in the direction of plate motion. Because of their geometry, age distributions, topographic expression and geochemical characteristics, most of these so-called hotspots are believed to mark the surface expression of upwelling mantle plumes^{1,7}.

Around 50 hotspots have been identified at Earth's surface^{8,9}. Of these, only ~20% occur on continents and, hence, most of our understanding of mantle plumes comes from hotspot tracks in oceanic settings. However, oceanic lithosphere is regularly recycled into the mantle through subduction, so if we are to understand plume-related volcanism before ~200 million years ago (Ma), which constitutes most of Earth's geological record¹⁰, we must learn: (1) how plumes interact with continental lithosphere; and (2) how this interaction affects the chemical composition and erupted volume of lavas at the surface. Here, we first combine observational constraints from surface geology,

geochronology and plate-motion histories to identify Earth's longest continental hotspot track in eastern Australia. We subsequently integrate constraints from seismology and geochemistry to ascertain how regional lithospheric thickness variations influence the volume and composition of plume-derived magmas along this track.

Cenozoic era volcanism in eastern Australia represents one of the world's most extensive intraplate volcanic regions¹¹ (Fig. 1). Three types of volcano are identified in the widely used classification of Wellman and McDougall¹¹, which we also adopt here: (1) central volcanoes, which are predominantly basaltic in composition but have felsic lava flows or intrusions, with lavas typically produced from central vents, often building large volcanic complexes; (2) lava fields, which are basaltic, extensive and thin and are often characterized by an abundance of small scoria, lava cones and maars; and (3) the leucite suite, which is dominated by low-volume, potassium-rich, leucite-bearing lavas. These volcanic classes are principally distinguished petrologically, with central volcanoes distinguished from lava-field volcanoes by the presence of felsic rock and both distinguished from the leucite suite by the absence of leucite¹². However, these classes also show considerable differences in age trends: ⁴⁰Ar–³⁹Ar and K–Ar geochronological studies demonstrate that both the central volcanoes and the leucite suite define age-progressive tracks that become younger to the south. The tracks identified, so far, include: (1) Comboyne—an ~770-km-long track that extends from Fraser Island in Queensland to Comboyne in New South Wales, displaying a record of volcanic activity from ~32 Ma to 16 Ma (ref. 13); (2) Canobolas—an ~760-km-long track, extending from Bunya in Queensland to Canobolas in New South Wales, recording volcanism from ~24 Ma to 12 Ma (ref. 14); (3) an ~400-km-long track, extending across central Queensland from Cape Hillsborough to Buckland and recording volcanism from ~34 Ma to 27 Ma (ref. 6); and (4) an ~650-km-long leucite-bearing track, extending from Bokhara River in New South Wales to Cosgrove in Victoria and displaying a record of volcanic activity from ~17 Ma to 9 Ma (ref. 5). These tracks are widely believed to mark the passage of mantle plumes beneath the northwards-migrating Australian plate^{5,6,13,14}. Lava-field volcanics, on the other hand, show no such age progression and are thought to be generated through an alternative process, with an edge-driven model being suggested for the formation of the Newer Volcanics Province (NVP)^{15,16}.

The central volcanoes of central Queensland and the leucite suite of New South Wales and Victoria have been considered unrelated^{5,6,11,14,17}, principally because the volcanic provinces identified in each differ dramatically in composition and eruptive volume, and they are separated by a volcanic gap of >650 km. However, their relative locations and ages suggest that they may be the surface expression of the same mantle plume, thus constituting a single hotspot track. In Fig. 1a we test this hypothesis by predicting volcanic locations along this track, using the reconstructed absolute motion of the Australian plate¹⁸. Specifically, we take the mapped locations of 15 volcanic centres (see Extended Data Table 1 and Extended Data Fig. 1),

¹Research School of Earth Sciences, The Australian National University, Canberra 2601, Australia. ²School of Geosciences, University of Aberdeen, Aberdeen AB243UE, UK. [†]Present address: Department of Geosciences & Natural Resource Management, University of Copenhagen, Copenhagen 1350, Denmark.

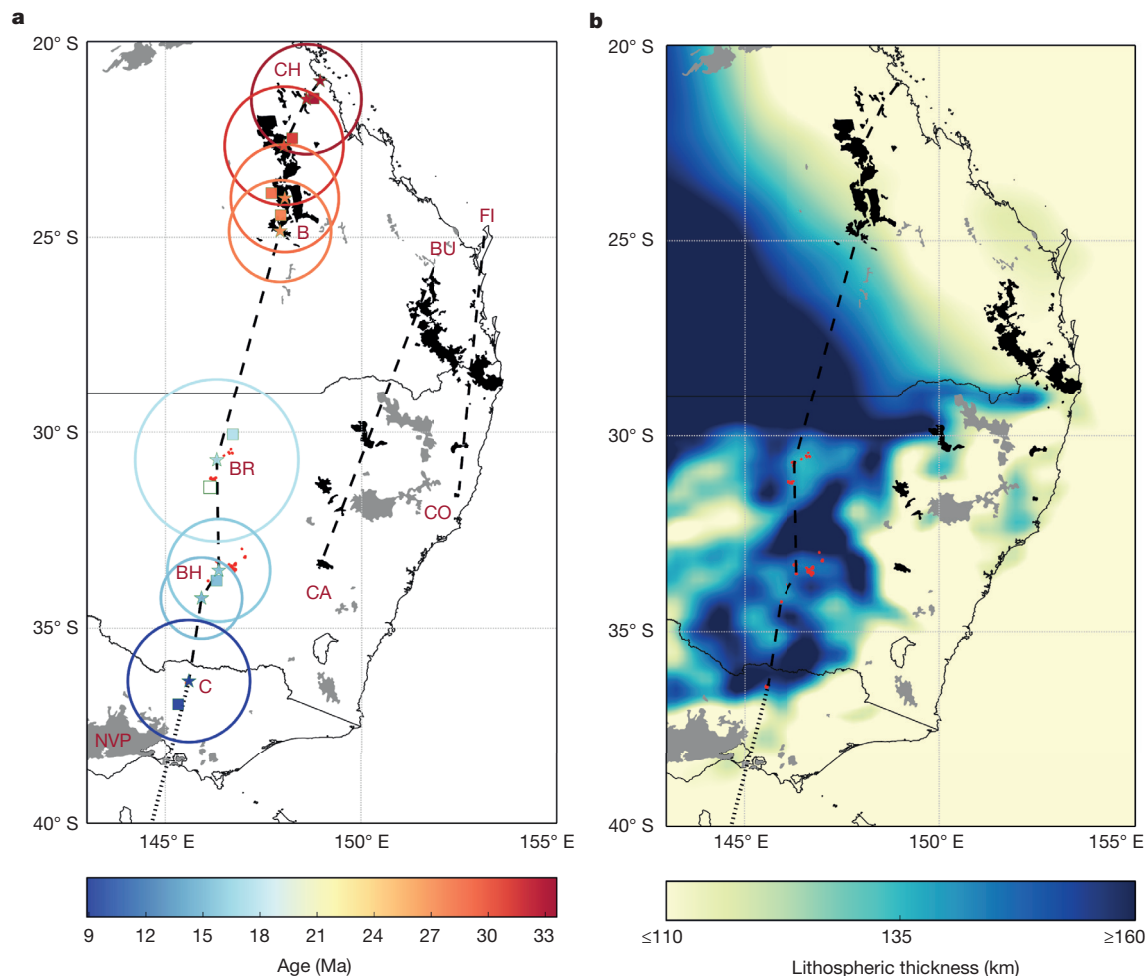


Figure 1 | The distribution and classification of eastern Australian Cenozoic volcanic centres and their relationship to regional lithospheric thickness variations. **a**, Volcanic centres, coloured following the classification of Wellman and McDougall¹¹, where black, grey and red denote central volcanoes, basaltic lava fields and leucite-bearing volcanics, respectively. Age-progressive hotspot tracks are denoted by black dashed lines. We identify Earth's longest continental hotspot track, the Cosgrove track, which extends across the Australian continent from Cape Hillsborough (CH) to Cosgrove (C). Stars denote a selection of volcanic centres along this track that are coloured by age (in Ma), while squares mark predicted volcanic locations, based on reconstructed absolute motions of the Australian plate¹⁸. If the predicted locations fall within

dated by ^{40}Ar – ^{39}Ar techniques^{5,6}, and predict their location at the time associated with the next dated volcanic centre (see Methods). Stars in Fig. 1a denote the locations of these dated volcanoes, while squares denote their predicted locations, based upon the reconstruction (note that for clarity, only 9 of the 15 dated volcanoes used in our analyses are shown: see Extended Data Fig. 2 for a comparable figure with all 15 volcanoes). Circles represent our estimate for the uncertainty in predicted locations, which arises through a combination of: (1) uncertainties in the underlying mantle plume diameter and the associated melt region's lateral extent^{3,19}; (2) the potential for plume drift^{8,20,21}; (3) the unpredictability introduced through preferential melt extraction pathways²²; and (4) uncertainties in the ^{40}Ar – ^{39}Ar ages of dated volcanic centres (see Methods and Extended Data Fig. 3). If a predicted location falls within these circles, then we consider that volcanic centre to be the surface expression of the same mantle plume (filled squares in Fig. 1a): 8 of the 9 volcanic centres shown in Fig. 1a, and 14 of the 15 volcanic centres in our full analyses, satisfy this criterion. The only volcanic centre that fails to satisfy this criterion is Begargo Hill of the leucite suite, which is consistently located further

the circles surrounding each dated volcanic centre (our measure of the uncertainty in predicted locations), then we consider that volcanic centre to be the surface expression of the same mantle plume: eight of nine volcanic centres satisfy this criteria (filled squares), while one does not (open square). The dotted line, to the south of Cosgrove, marks the predicted extension of the Cosgrove hotspot track, towards its present-day location, which is illustrated in Extended Data Fig. 2. B, Buckland; BH, Begargo Hill; BR, Bokhara River; BU, Bunya; CA, Canobolas; CO, Comboyne; FI, Fraser Island; NVP, Newer Volcanics Province. **b**, The same volcanic centres, plotted above an estimate of lithospheric thickness, highlighting a clear correlation between lithospheric thickness and volcanic outcrop and composition along the Cosgrove hotspot track.

south than is predicted by our reconstruction. We speculate that this points towards either a rapid phase of southwards plume motion (exceeding 4 cm yr^{-1}) from $\sim 17\text{ Ma}$ to 15 Ma , a change in the melt extraction pathway, or a combination of both. In support of these ideas, we note that: (1) variable plume migration rates have been observed elsewhere²⁰ and are also predicted in global mantle convection simulations^{8,21}; and (2) Begargo Hill lies to the south of a region of increased lithospheric thickness (discussed later), which may focus sub-lithospheric plume material and any associated melt southwards, as a result of Australia's rapid motion towards the north. Most notably, however, the northernmost dated leucite-bearing volcano satisfies the aforementioned location criterion, confirming that the central volcanoes of central Queensland and the leucite suite of New South Wales and Victoria are the surface expression of the same mantle plume. Together, they constitute Earth's longest continental hotspot track, which we call the Cosgrove track. This knowledge, however, leads to further questions. Specifically, since these volcanic centres are the surface expression of the same mantle plume, why does the Cosgrove track display large volcanic gaps? What drives the consid-

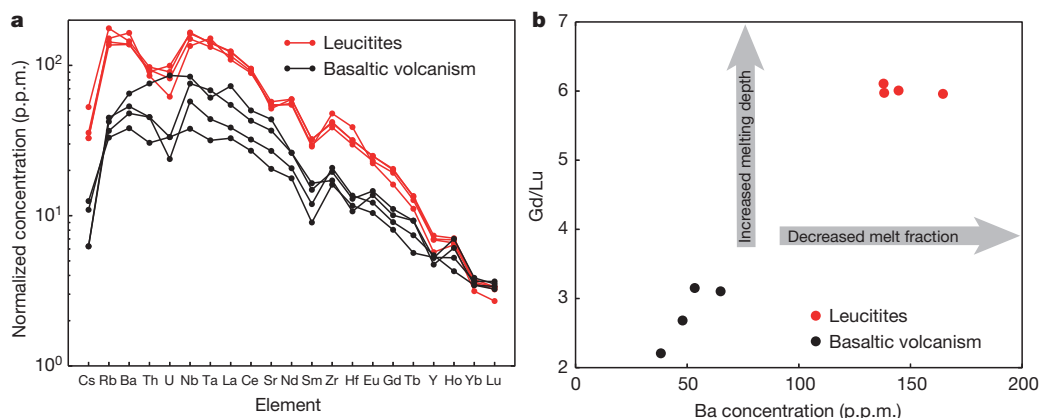


Figure 2 | Trace-element abundances of volcanic samples along the Cosgrove hotspot track. **a**, Trace-element concentrations (in p.p.m.), normalized to primitive mantle values³¹, from published analyses of four leucitite-bearing and four mantle-derived basaltic samples along the southern

and northern sections of the Cosgrove track, respectively (see Methods). **b**, A plot of the Gd/Lu ratio (a proxy for melting depth) against the barium (Ba) concentration (a tracer for the extent of melting) for these samples.

erable variations in volume and chemical composition of plume-derived magmas between the central volcanoes and the leucitite suite? Are both of these characteristics related? To answer these fundamental questions, we combine observational constraints from seismology and geochemistry.

We first generate a map of lithospheric thickness, by combining constraints from recent three-dimensional body-wave tomography results^{15,23} with the regional Australian Seismological Reference Model (AuSREM)²⁴ (Fig. 1b) (for further details, including uncertainties, see Methods and Extended Data Figs 4–6). The main features evident in Fig. 1b include: (1) the contrast between thicker lithosphere in the centre of Australia and thinner lithosphere to the east, which is consistent with a transition from Precambrian central to Phanerozoic eastern Australia and then oceanic lithosphere outboard of the continental margin; (2) a zone of thin lithosphere, bound to the east by a zone of intermediate-thickness lithosphere of similar width, which extends southwestwards from $\sim 30^\circ$ S through central New South Wales into northern Victoria¹⁵; and (3) considerable changes in lithospheric thickness over relatively short horizontal distances. It is generally accepted that these lithospheric ‘steps’ will produce complex flows²⁵ and, as noted previously, this led to the suggestion of an edge-related model for the formation of the NVP^{15,16,25}. Indeed, such edge-related mechanisms are probably also applicable to other lava-field volcanism in the region: as illustrated in Fig. 1b, all lava-field volcanic provinces lie adjacent to substantial steps in lithospheric thickness, above comparatively thin lithosphere, thus providing a favourable setting^{15,16}.

What remains poorly understood, however, is how mantle plumes interact with these lithospheric thickness variations and, specifically, how they influence the volume and composition of plume-derived magmas. Intriguing trends are evident along the Cosgrove track in Fig. 1b: (1) volcanic gaps occur in regions where lithospheric thickness exceeds ~ 150 km; (2) the basaltic and felsic central volcanoes of central Queensland occur in regions where lithospheric thickness is less than ~ 110 km; and (3) low-volume, leucitite-bearing volcanism to the south occurs, exclusively, in regions of intermediate lithospheric thickness, with volcanic gaps within the leucitite suite also coinciding with regions of thicker lithosphere. These unambiguous trends suggest that the thickness of overlying lithosphere is dictating the volume and composition of plume-derived magmas, by limiting the rise height of the underlying plume and, hence, the degree of partial melting. We infer that the underlying mantle plume: (1) cannot rise to shallow enough depths to induce decompression melting in regions where lithospheric thickness exceeds ~ 150 km, thus providing an explanation for the volcanic gaps along the Cosgrove track and placing the first observational constraint on the maximum melting depth of mantle

plumes beneath continents (excluding ultra-volatile melts that form kimberlites and carbonatites); (2) undergoes high-degree partial melting beneath comparatively thin lithosphere to produce basaltic and felsic central volcanoes along the northern segment of the Cosgrove track; and (3) undergoes very low-degree partial melting in regions of intermediate lithospheric thickness, thus facilitating the production of low-volume leucitite-bearing volcanics towards the southern end of the Cosgrove track.

To ascertain whether or not these inferences are compatible with geochemical observations, we collated previously published trace-element data from volcanic outcrops along the Cosgrove track^{17,26} (see Methods and Extended Data Tables 2, 3). Although this data set is limited, consisting of only eight data points (four from the central volcanoes of central Queensland and four from the leucitite suite), it provides a basis for testing our hypothesis. As mantle material undergoes partial melting, incompatible trace elements, such as barium (Ba), are preferentially transferred into the melt. Indeed, for the most incompatible elements, their concentration in the melt is, to a good approximation, inversely proportional to the melt fraction. As the degree of partial melting increases, their concentrations are subsequently diluted²⁷. Accordingly, if the available trace-element data supported the inferences presented earlier, leucitite-bearing volcanics would display higher concentrations of incompatible trace elements when compared to the basaltic central volcano samples. Such trends are indeed apparent in the trace-element concentrations plotted in Fig. 2: leucitite samples display barium concentrations that exceed those of basaltic samples by a factor of ~ 3 . Our inference that the melt fraction along the Cosgrove track is controlled by lithospheric thickness, which limits the rise height of plumes, should also leave a discernible signature in the trace-element concentrations illustrated in Fig. 2. Melting at greater depth, in the presence of higher proportions of garnet²⁸, will sequester the heavy rare-earth elements, such as lutetium (Lu), with respect to the middle rare-earth elements, such as gadolinium (Gd). Consequently, samples from the leucitite suite should display elevated Gd/Lu ratios when compared to the basaltic samples from central Queensland. As illustrated in Fig. 2, this is the case: leucitites typically possess Gd/Lu ratios of ~ 6 , compared to ~ 2 – 3 for the basalts of central Queensland. Our hypothesis, therefore, is supported by the available trace-element data.

An aspect of the Cosgrove track that has not been addressed is why basaltic and felsic central volcanoes do not re-emerge to the south of Cosgrove, in a region of comparatively thin lithosphere. As mantle plumes have finite lifetimes^{8,21}, it is possible that the underlying plume faded at ~ 8 Ma, although this would be an unlikely coincidence. We speculate that an alternative mechanism is at play: it has previously

been demonstrated¹⁵ that regional three-dimensional lithospheric thickness variations, coupled with rapid northwards motion of the Australian plate, gives rise to a focused edge-driven convection (EDC) cell to the west of the Cosgrove track, in the vicinity of the NVP, thus providing a mechanism for the localization of lava-field volcanism to this region. However, an explanation for the onset of NVP volcanism, at ~5 Ma (ref. 11), has remained elusive: the lithospheric thickness variations driving EDC in this region are probably long-lived²⁹, which is difficult to reconcile with comparatively recent volcanism. Our reconstructions, however, place the mantle plume that generated the Cosgrove hotspot track <50 km to the east of the NVP, from ~6.5 Ma to 5 Ma. We speculate that the capture and entrainment of this plume, into a pre-existing EDC cell, was the trigger for magmatism within the NVP and explains the absence of a hotspot track to the south of Cosgrove. In support of these ideas, we note that although EDC is expected to occur on all lithospheric steps, including those to the east of the plume's predicted passage, the dominant cell in this region lies directly beneath the NVP¹⁵. Accordingly, preferential westwards flow (and, hence, entrainment of plume material) into the NVP region, is to be expected, and was indeed evident in our previous models¹⁵. To our knowledge, the interaction between mantle plumes and EDC has not been documented elsewhere. However, this process: (1) has important implications for the surface expression of mantle plumes in the vicinity of step changes in lithospheric thickness; and (2) provides a solution to the global puzzle of why step changes in lithospheric thickness, which occur along craton edges and at passive margins, produce volcanism only at isolated locations, thus complementing our previous study¹⁵. Finally, we note that the predicted present-day location of the mantle plume that generated the Cosgrove hotspot track lies to the northwest of Tasmania, which coincides with a region of recent seismicity and is at the western limit of the so-called East Australia Plume System imaged previously using finite frequency tomography³⁰ (see Extended Data Fig. 2).

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 14 May; accepted 7 July 2015.

Published online 14 September 2015.

1. Morgan, W. J. Convection plumes in the lower mantle. *Nature* **230**, 42–43 (1971).
2. Davies, G. F. Thermomechanical erosion of the lithosphere by mantle plumes. *J. Geophys. Res.* **99**, 15709–15722 (1994).
3. Farnetani, C. G. & Richards, M. A. Thermal entrainment and melting in mantle plumes. *Earth Planet. Sci. Lett.* **136**, 251–267 (1995).
4. White, R. S. & McKenzie, D. Mantle plumes and flood basalts. *J. Geophys. Res.* **100**, 17543–17585 (1995).
5. Cohen, B. E., Knesel, K. M., Vasconcelos, P. M., Thiede, D. S. & Hergt, J. M. ⁴⁰Ar/³⁹Ar constraints on the timing and origin of Miocene leucite volcanism in southeastern Australia. *Aust. J. Earth Sci.* **55**, 407–418 (2008).
6. Cohen, B. E., Knesel, K. M., Vasconcelos, P. M. & Schellart, W. P. Tracking the Australian plate motion through the Cenozoic: constraints from ⁴⁰Ar/³⁹Ar geochronology. *Tectonics* **32**, 1371–1383 (2013).
7. Duncan, R. A. & Richards, M. A. Hotspots, mantle plumes, flood basalts and true polar wander. *Rev. Geophys.* **29**, 31–50 (1991).
8. Steinberger, B. Plumes in a convecting mantle: models and observations for individual hotspots. *J. Geophys. Res.* **105**, 11127–11152 (2000).
9. Courtillot, V., Davaille, A., Besse, J. & Stock, J. Three distinct types of hotspots in the Earth's mantle. *Earth Planet. Sci. Lett.* **205**, 295–308 (2003).

10. Campbell, I. H. & Griffiths, R. W. The changing nature of mantle hotspots through time: implications for the geochemical evolution of the mantle. *J. Geol.* **100**, 497–523 (1992).
11. Wellman, P. & McDougall, I. Cainozoic igneous activity in eastern Australia. *Tectonophysics* **23**, 49–65 (1974).
12. Johnson, R. W. (ed.) *Intraplate Volcanism in Eastern Australia and New Zealand* Ch. 1.1 (Cambridge Univ. Press, 1989).
13. Knesel, K. M., Cohen, B. E., Vasconcelos, P. M. & Thiede, D. S. Rapid change in drift of the Australian plate records collision with Ontong Java plateau. *Nature* **454**, 754–757 (2008).
14. Sutherland, F. L., Graham, I. T., Meffre, S., Zwingmann, H. & Pogson, R. E. Passive-margin prolonged volcanism, east Australian plate: outbursts, progressions, plate controls and suggested causes. *Aust. J. Earth Sci.* **59**, 983–1005 (2012).
15. Davies, D. R. & Rawlinson, N. On the origin of recent intra-plate volcanism in Australia. *Geology* **42**, 1031–1034 (2014).
16. King, S. D. & Anderson, D. L. Edge-driven convection. *Earth Planet. Sci. Lett.* **160**, 289–296 (1998).
17. Ewart, A., Chappell, B. W. & Menzies, M. A. An overview of the geochemical and isotopic characteristics of the eastern Australian Cainozoic volcanic provinces. *J. Petrol.* **1**, 225–273 (1988).
18. Torsvik, T. H., Steinberger, B., Gurnis, M. & Gaina, C. Plate tectonics and net lithosphere rotation over the past 150 Myr. *Earth Planet. Sci. Lett.* **291**, 106–112 (2010).
19. Leitch, A. M. & Davies, G. F. Mantle plumes and flood basalts: enhanced melting from plume ascent and an eclogite component. *J. Geophys. Res.* **106**, 2047–2059 (2001).
20. Tarduno, J. A. *et al.* The Emperor Seamounts: southward motion of the Hawaiian hotspot plume in Earth's mantle. *Science* **301**, 1064–1069 (2003).
21. Davies, D. R. & Davies, J. H. Thermally-driven mantle plumes reconcile multiple hotspot observations. *Earth Planet. Sci. Lett.* **278**, 50–54 (2009).
22. Sleep, N. H. Lateral flow of hot plume material ponded at sublithospheric depths. *J. Geophys. Res.* **101**, 28065–28083 (1996).
23. Rawlinson, N., Kennett, B. L. N., Salmon, M. & Glen, R. A. in *The Earth's Heterogeneous Mantle: A Geophysical, Geodynamical, and Geochemical Perspective* Ch. 2 (eds Khan, A., & Deschamps, F.) 47–78 (Springer, 2015).
24. Kennett, B. L. N., Fichtner, A., Fishwick, S. & Yoshizawa, K. Australian Seismological Reference Model (AuSREM): mantle component. *Geophys. J. Int.* **192**, 871–887 (2013).
25. Farrington, R. J., Stegman, D. R., Moresi, L. N., Sandiford, M. & May, D. A. Interactions of 3D mantle flow and continental lithosphere near passive margins. *Tectonophysics* **483**, 20–28 (2010).
26. Paul, B., Hergt, J. M. & Woodhead, J. D. Mantle heterogeneity beneath the Cenozoic volcanic provinces of central Victoria inferred from trace-element and Sr, Nd, Pb and Hf isotope data. *Aust. J. Earth Sci.* **52**, 243–260 (2005).
27. Hofmann, A. W. in *Treatise on Geochemistry* Vol. 2 (ed. Carlson, R. W.) 61–101 (Elsevier, 2003).
28. Ringwood, A. E. *Composition and Petrology of the Earth's Mantle* (McGraw-Hill, 1975).
29. Rawlinson, N. *et al.* Complex continental growth along the proto-Pacific margin of East Gondwana. *Geology* **42**, 783–786 (2014).
30. Montelli, R., Nolet, G., Dahlen, F. A. & Masters, G. A catalogue of deep mantle plumes: new results from finite-frequency tomography. *Geochim. Geophys. Geosys.* **7**, Q11007 (2006).
31. McDonough, W. F. & Sun, S.-S. The composition of the Earth. *Chem. Geol.* **120**, 223–253 (1995).

Acknowledgements D.R.D. is funded by an Australian Research Council Future Fellowship (FT140101262). G.I. acknowledges support from the Ringwood Fellowship at the Australian National University. Digital geological data were provided by Geosciences Australia.

Author Contributions D.R.D. conceived this study and integrated all interdisciplinary observational constraints. N.R. created the lithospheric thickness map by combining constraints from the AuSREM reference model and body-wave data from the WOMBAT array. He also devised and implemented the method for estimating uncertainty in lithospheric thickness. G.I. performed the hotspot-track reconstruction and estimated the associated uncertainties. D.R.D. and I.H.C. undertook the geochemical synthesis. D.R.D. wrote the paper, following discussion with, and contributions from, all authors.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to D.R.D. (rhodri.davies@anu.edu.au).

METHODS

⁴⁰Ar–³⁹Ar dated volcanic centres. Of the volcanic centres that make up the Cosgrove hotspot track, 15 have recently been dated via ⁴⁰Ar–³⁹Ar geochronology. These include nine volcanic centres from the central volcanoes of Central Queensland⁶ and six volcanic centres from the leucitite suite of New South Wales and Victoria⁵. These data are summarized in Extended Data Table 1, with localities also illustrated in Extended Data Fig. 1.

Reconstructed hotspot track. In Fig. 1a, we test the hypothesis that the central volcanoes of central Queensland and the leucitite suite of New South Wales and Victoria represent the surface expression of the same mantle plume, thus constituting a single hotspot track, by predicting volcanic locations along this track, using the reconstructed absolute motion of the Australian plate¹⁸. Specifically, we take the mapped location of 15 volcanic centres^{5,6} (Extended Data Table 1) and predict their location at the time associated with the next (in a temporal sense) dated volcanic centre. Stars in Fig. 1a denote the locations of these dated volcanoes, while squares denote their predicted locations, based upon the reconstruction. Note that for illustrative purposes, only 9 of the 15 dated volcanoes used in our analyses were shown in Fig. 1a. A comparable figure, for all 15 dated volcanoes, is shown in Extended Data Fig. 2.

In doing this, we have implicitly assumed that the underlying mantle plume is a fixed point source at the base of the lithosphere and that melt transport towards Earth's surface is not influenced by pre-existing lithospheric structure. However, palaeomagnetic data indicate that mantle plumes drift slowly, at a fraction of surface plate velocities^{20,32,33}, an inference that is supported by global mantle convection simulations^{8,21}. In addition, plume tails have an upper-mantle diameter that probably exceeds 200 km and produce a sub-lithospheric melt region that may be up to half of this size^{3,19}, while lithospheric structure has an important role in controlling the lateral transport of plume material and melt-extraction pathways²². We therefore estimate the uncertainty in our predicted locations (Fig. 1a, circles, and Extended Data Fig. 2), by allowing for plume motion at a modest rate of up to 1 cm yr⁻¹, and adding 100 km to account for the melt region's lateral extent and the unpredictability introduced through preferential melt extraction pathways. Furthermore, we propagate the uncertainty on the measured ages of volcanic samples onto the predicted volcanic locations^{5,6}. Specifically, we use the measured age uncertainties^{5,6} to shorten/lengthen the duration of the temporal stages over which the instantaneous angular velocities/poles¹⁸ are applied. This results in two different predictions of geographic positions. We take the geodesic distance between these two as a representative value of the uncertainty contribution, on the predicted locations, which is associated with age. If a predicted location falls within the resulting uncertainty circles, we consider that volcano to be the surface expression of the same mantle plume (Fig. 1a, filled squares, and Extended Data Fig. 2): 8 of the 9 volcanic centres shown in Fig. 1a and 14 of the 15 shown in Extended Data Fig. 2 satisfy this criterion. The sensitivity of our results to the uncertainty parameters specified (that is, the rate of plume motion and the melt zone's lateral extent) is illustrated in Extended Data Fig. 3.

Lithospheric thickness estimate. The lithospheric thickness estimate given in this study is derived from recent three-dimensional seismic tomography results^{15,23}, which exploit body-wave arrival time information from the WOMBAT transportable array in eastern Australia. WOMBAT is the largest experiment of its type in the southern hemisphere, which, to date, comprises a cumulative total of ~700 stations, with spacings varying from 50 km on the mainland to 15 km in Tasmania. Extended Data Figure 4 shows the location of the 12 sub-arrays used in this study. Note that although two additional arrays have been deployed to the northeast, the associated data are not yet available for use.

FMTOMO^{34,35} is used to invert relative arrival time residuals of various global P-wave phases for three-dimensional velocity variations. Owing to the dominance of short-period seismometers (1 Hz natural frequency) in WOMBAT, S-wave arrivals are difficult to identify, which is why we favour P-wave arrivals. Two limitations of using relative P-wave arrival times in seismic tomography are that absolute velocities are not constrained, and the long-wavelength structure (greater than the aperture of the sub-arrays) is filtered out during the inversion. We overcome these limitations by using the AuSREM mantle model²⁴ as a starting model in the inversion, which has a horizontal resolution of ~200–250 km. Assuming that AuSREM is accurate at this structural wavelength, the problem of suturing together relative arrival time data from different arrays is overcome³⁶ and both the absolute and relative velocities can be considered (although the latter will be more reliable).

The use of three-dimensional mantle velocity models to derive estimates of lithospheric thickness is common. In Australia, a number of different estimates have been made based upon S-wave velocity models obtained from surface wave tomography^{37–39}. Proxies for the depth to the base of the lithosphere include a

particular choice of velocity contour, a decrease in velocity within a certain depth range, and the behaviour of the velocity gradient. In our case, we have a three-dimensional P-wave velocity model rather than a three-dimensional S-wave velocity model, which is likely to be less sensitive to the presence of a mechanically weak sub-lithosphere mantle layer. Moreover, the limited depth resolution caused by the sub-vertical nature of the incident paths means that using the depth to the base of fast velocity anomalies, or some chosen contour, is unlikely to yield robust results, particularly in terms of absolute values. Extended Data Figure 5 shows a slice through the P-wave velocity model at 120 km depth, which reveals a clear division between higher velocities to the west and lower velocities to the east, which is consistent with a transition from Precambrian central and western Australia, to Phanerozoic eastern Australia and finally oceanic lithosphere outboard of the continental margin. Regions of elevated velocities probably correspond to thicker lithosphere, as the pattern of velocity variations shares a first-order resemblance to the lithospheric thickness estimates of previous studies^{39,40}.

The AuSREM model⁴¹ also includes a lithospheric thickness estimate, based on velocity gradients from surface and body-wave tomography and refracted waves in the mantle. This estimate shows a broad consistency with other studies^{39,40}, with the main difference being in the absolute depths to the base of the lithosphere rather than the pattern of depth variations. Minimum depths vary between ~50 km and 70 km in the vicinity of the eastern seaboard, whereas maximum depths vary between 180 km and 220 km in the central western section of our model region. On the basis of changes in the vertical gradient of SV-wave velocity, one study³⁹ provides upper and lower bounds on the transition from lithosphere to sub-lithospheric mantle. This transition exceeds 50 km at several locations. Both the models described in refs 40 and 41 correspond more closely to the lower bound estimates of ref. 39.

Owing to the potential for vertical smearing in the three-dimensional WOMBAT model, we instead take the vertically averaged velocity over a prescribed depth range and calibrate this against minimum and maximum regional lithospheric thickness estimates from previous studies. Thus, the maximum average velocity will equate to the thickest lithosphere, while the minimum average velocity will equate to the thinnest lithosphere, with a linear relationship assumed between depth and average velocity. This approach essentially assumes that variations in the thickness of a higher-velocity lithosphere above a lower velocity sub-lithospheric mantle is responsible for the observed variations in arrival time. Although this is a relatively crude assumption, if we apply this scheme to the P-wave component of the AuSREM model over a depth range of 50–200 km, then the resultant estimate of lithospheric thickness is similar to the AuSREM lithosphere thickness estimate. Extended Data Figure 6a shows the result of using the same 50–200 km depth range with the three-dimensional WOMBAT tomography model (equivalent to Fig. 1), noting that north of ~28° S, we revert to the AuSREM model, as there is no additional data coverage in this region. The pattern of depth variations is remarkably similar to the pattern of velocity variations at 120 km depth in Extended Data Fig. 5, demonstrating that velocities do not vary significantly with depth in the uppermost mantle. Compared to previous lithospheric thickness estimates for the Australian continent, the main difference is that the new model includes a zone of (on average) intermediate thickness lithosphere that extends southwards from about 30° S through central New South Wales before terminating in northern Victoria. This feature is crucial to the results of this study, as it hosts the leucitite-bearing volcanics.

To investigate the robustness of our lithospheric thickness estimate, we perform the same calculation but vary the input parameters. In particular, we vary the minimum thickness of the model between 50 km and 70 km, and the maximum thickness between 180 km and 220 km, which reflects the range of extremes in our geographic region, between the models of refs 39, 40 and 41. Moreover, in an attempt to account for limited vertical resolution in the mantle and the potential for unresolved crustal structure to smear into mantle structure, we vary the minimum depth over which we take the velocity average between 50 km and 100 km and the maximum depth from 180 km to 230 km. In the former case, the three-dimensional tomographic model explicitly includes the AuSREM crustal and Moho depth model⁴¹ to minimize this trade-off effect, so this range is probably quite generous. However, with a minimum depth set at 100 km, we are likely to ignore shallower anomalies that are unrelated to lithospheric thickness. The range of maximum depths was chosen to account for smearing effects at depth; for example, if high velocity lithosphere terminating at 200 km depth is recovered as slightly lower velocity lithosphere terminating at 230 km depth owing to smearing, then it may manifest as thicker lithosphere if the vertical averaging is taken to a depth of 230 km rather than 200 km. We generate a total of 540 models by iterating over these four depth ranges, using an increment of 10 km. Extended Data Figure 6b shows the standard deviation of this model ensemble. In general, the uncertainty is dominated by our chosen range of minimum and maximum

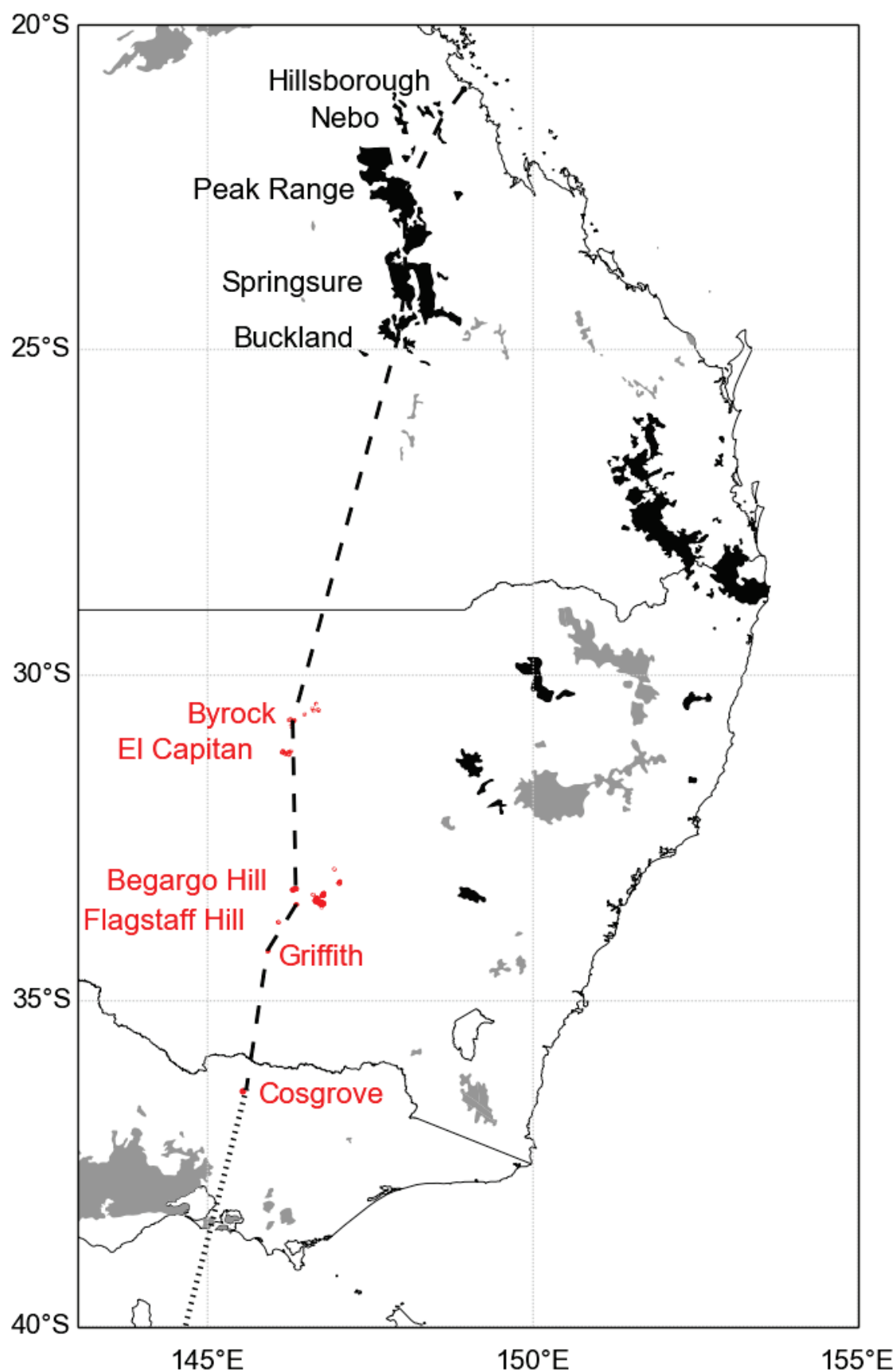
lithosphere thicknesses, but there is also a contribution from our choice of depth range over which to vertically average the velocity. The end result indicates that, in most regions of the lithosphere thickness model, the standard deviation of the uncertainty is ~ 10 km or less, implying that the first-order features of Extended Data Fig. 6a (and Fig. 1b) are probably correct. Moreover, the pattern of thickness variations is more robust than the absolute variations, which are dependent on the minimum and maximum thickness bounds that we obtain from previously published models.

Geochemical samples. Leucitites of New South Wales and Victoria: trace-element data from published analyses of leucite samples along the Cosgrove track are summarized in Extended Data Tables 2 and 3. These include three samples from Cosgrove in Victoria²⁶ (acquired by solution ICPMS) and one sample from Condobolin in New South Wales¹⁷ (acquired by XRF). Limited trace-element data for three samples at Begargo Hill, Flagstaff and El Capitan in New South Wales, from ref. 42 (acquired by isotope dilution), are also included. Given that trace-element concentrations are only provided for four elements (Rb, Sr, Nd and Sm) in ref. 42, they are not plotted in Fig. 2. Nonetheless, the limited trace-element concentrations provided are consistent with more complete data sets^{17,26}.

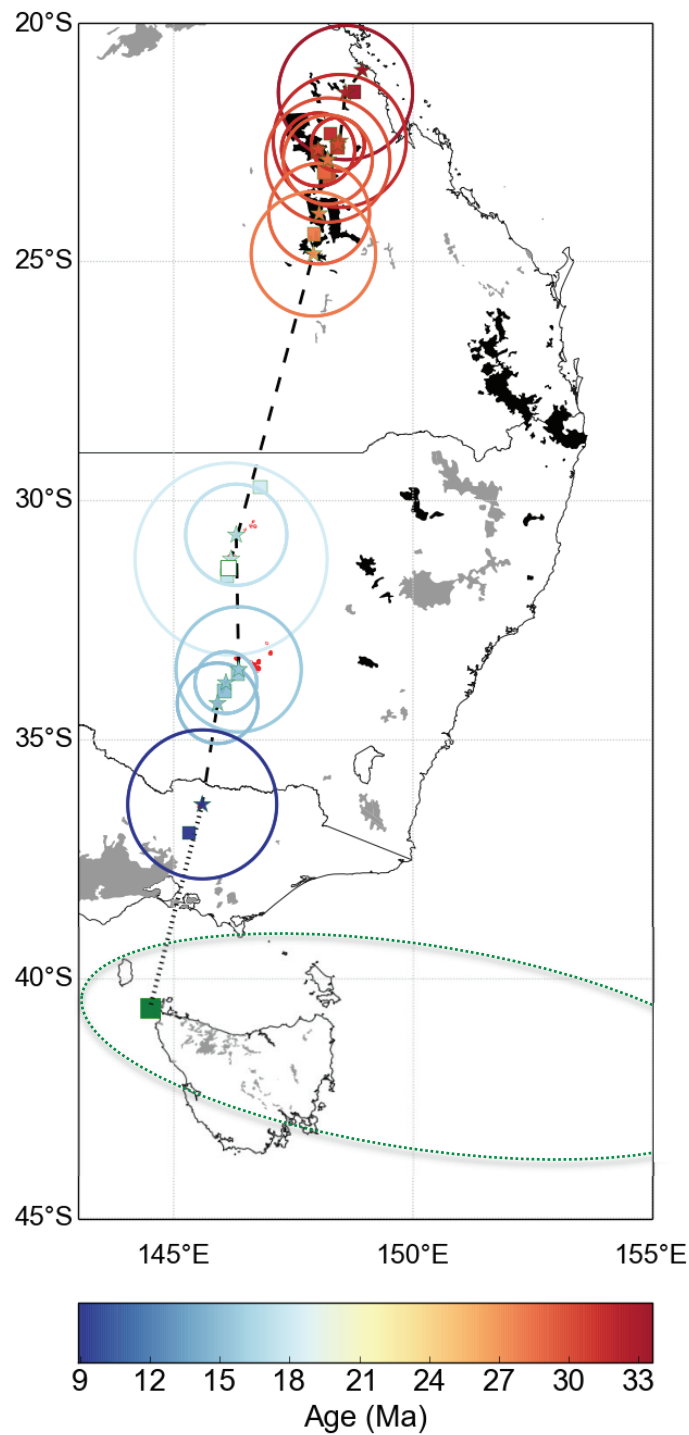
Basaltic samples from the central volcanoes of central Queensland: we restrict our choice of trace-element data to samples with a composition that is likely to reflect their mantle source (that is, $\text{SiO}_2 < 50\%$; $\text{MgO} > 9\%$). The selected samples also have no negative Ni/Ta or Eu anomalies, which indicates no detectable crustal contamination or fractional crystallization within the crust.

Geochemical data, with the exception of ref. 17, can be downloaded from the GEOROC database at EarthChem (<http://www.earthchem.org/>).

32. Antretter, M., Steinberger, B., Heider, F. & Soffel, H. Paleolatitudes of the Kerguelen hotspot: new paleomagnetic results and dynamic modelling. *Earth Planet. Sci. Lett.* **203**, 635–650 (2002).
33. Tarduno, J. A., Bunge, H.-P., Sleep, N. & Hansen, U. The bent Hawaiian-Emperor hotspot track: inheriting the mantle wind. *Science* **324**, 50–53 (2009).
34. Rawlinson, N. & Urvoy, M. Simultaneous inversion of active and passive source datasets for 3-D seismic structure with application to Tasmania. *Geophys. Res. Lett.* **33**, L24313 (2006).
35. Rawlinson, N., Tkalcic, H. & Reading, A. M. Structure of the Tasmanian lithosphere from 3D seismic tomography. *Aust. J. Earth Sci.* **57**, 381–394 (2010).
36. Rawlinson, N., Salmon, M. & Kennett, B. L. N. Transportable seismic array tomography in southeast Australia: illuminating the transition from Proterozoic to Phanerozoic lithosphere. *Lithos* **189**, 65–76 (2014).
37. Simons, F. J. & van der Hilst, R. D. Age-dependent seismic thickness and mechanical strength of the Australian lithosphere. *Geophys. Res. Lett.* **29**, 1529 (2002).
38. Fishwick, S. & Rawlinson, N. 3-D structure of the Australian lithosphere from evolving seismic datasets. *Aust. J. Earth Sci.* **59**, 809–826 (2012).
39. Yoshizawa, K. Radially anisotropic 3-D shear wave structure of the Australian lithosphere and asthenosphere from multi-mode surface waves. *Phys. Earth Planet. Inter.* **235**, 33–48 (2014).
40. Fishwick, S., Heintz, M., Kennett, B. L. N., Reading, A. M. & Yoshizawa, K. Steps in lithospheric thickness within eastern Australia: evidence from surface wave tomography. *Tectonics* **27**, TC4009 (2008).
41. Kennett, B. L. N. & Salmon, M. AuSREM: Australian seismological reference model. *Aust. J. Earth Sci.* **59**, 1091–1103 (2012).
42. Nelson, D. R., McCulloch, M. T. & Sun, S.-S. The origins of ultrapotassic rocks as inferred from Sr, Nd, and Pb isotopes. *Geochim. Cosmochim. Acta* **50**, 231–245 (1986).

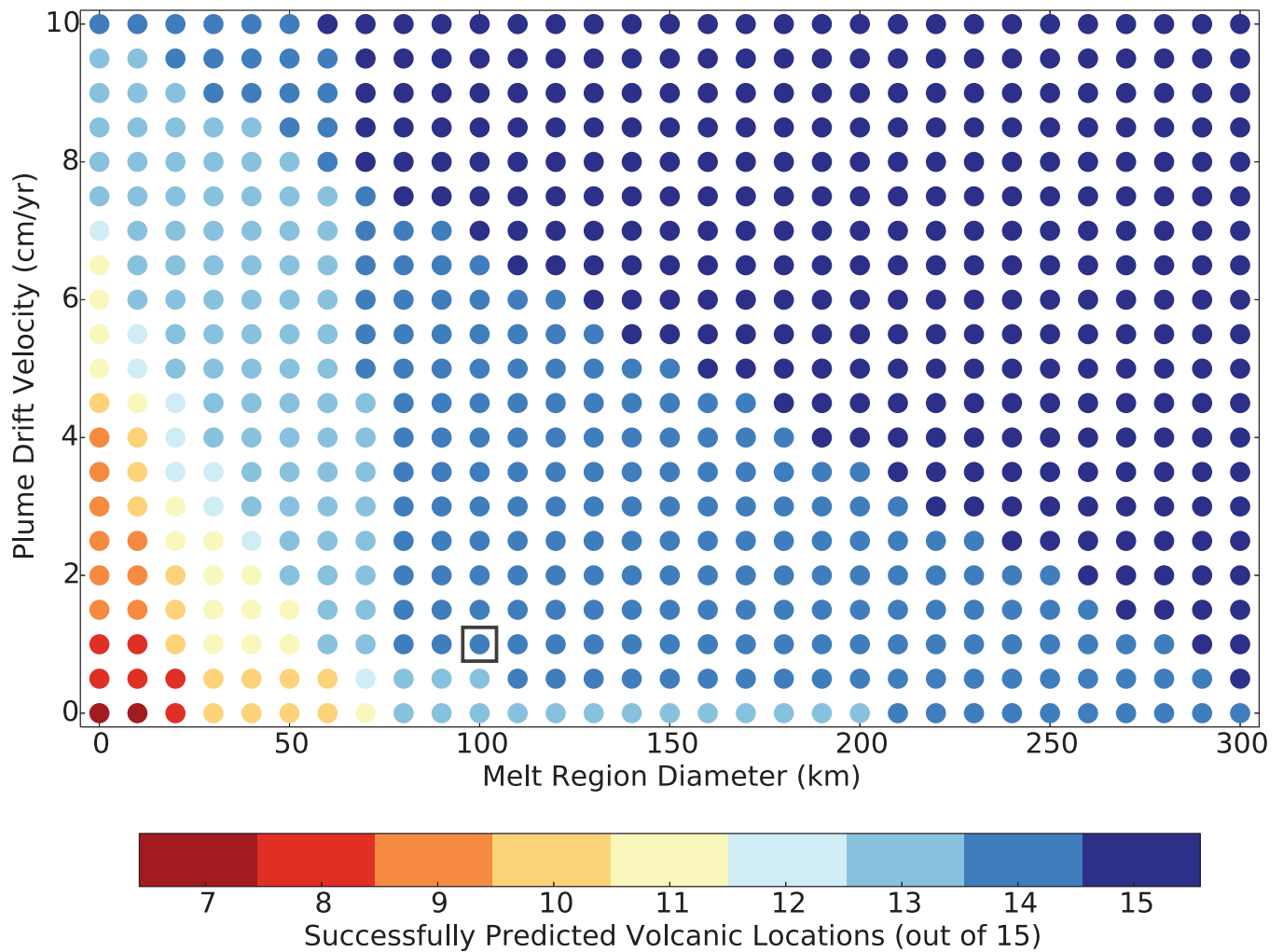


Extended Data Figure 1 | Locations of the 15 volcanic centres used in our reconstruction of the Cosgrove hotspot track. The hotspot track is indicated with a dashed line.



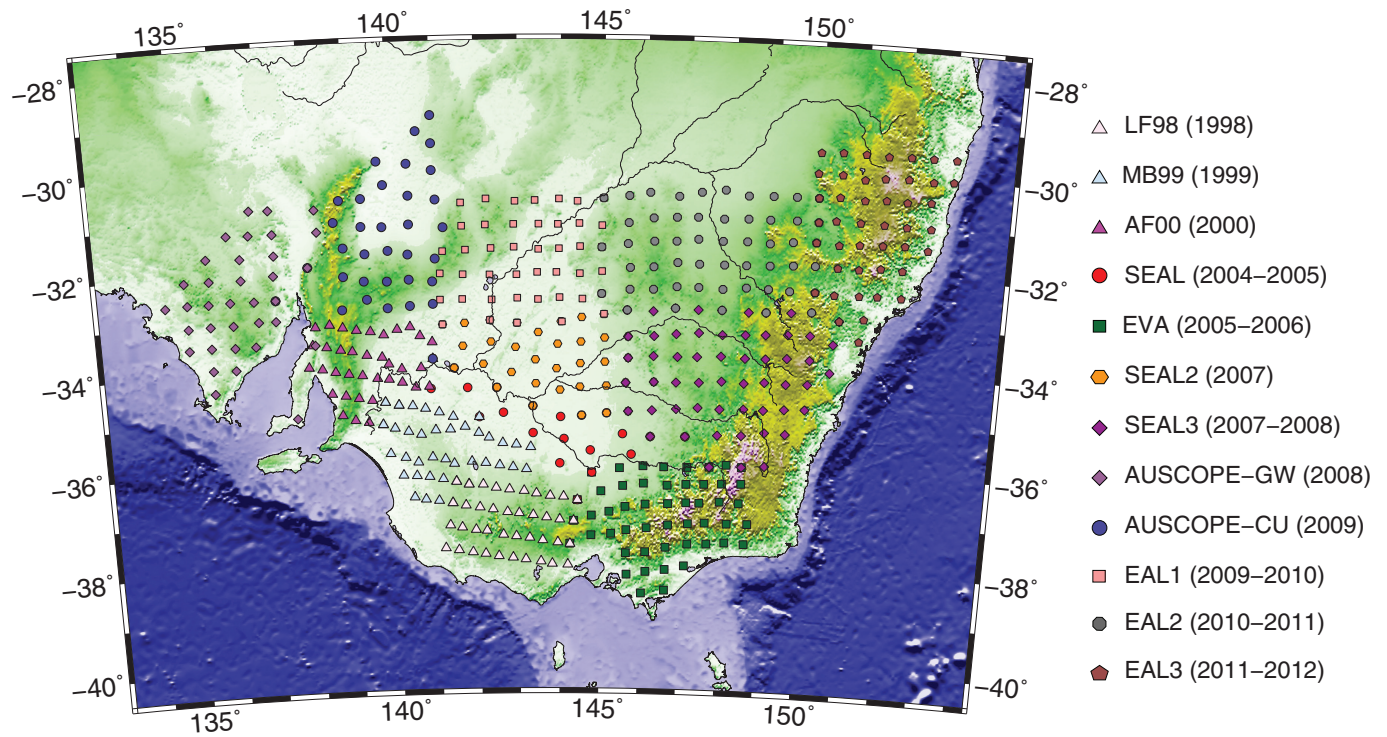
Extended Data Figure 2 | The Cosgrove hotspot track. As in Fig. 1a but incorporating all 15 dated volcanic complexes and extended southwards to show the predicted present-day location of the underlying mantle plume

(green square to the northwest of Tasmania). The approximate location of the East Australia Plume System, imaged previously using finite frequency tomography³⁰, is marked by the dotted green line.

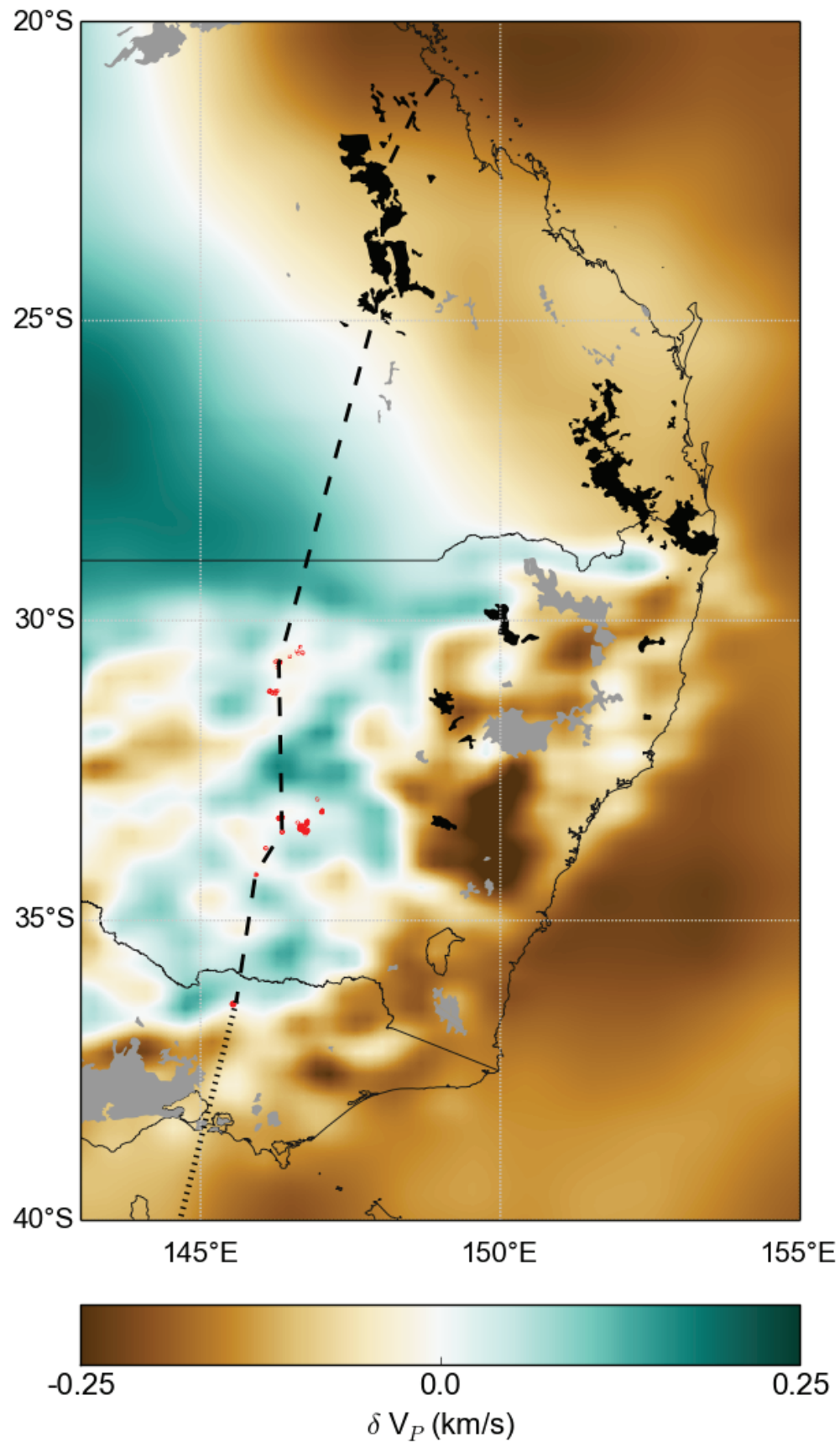


Extended Data Figure 3 | Reconstruction score map. The number of predicted volcanic centre locations, from a total of 15 (listed in Extended Data Table 1), that fall within the uncertainty circles surrounding the dated volcanic centres, for a range of plume drift velocities and melt region diameters.

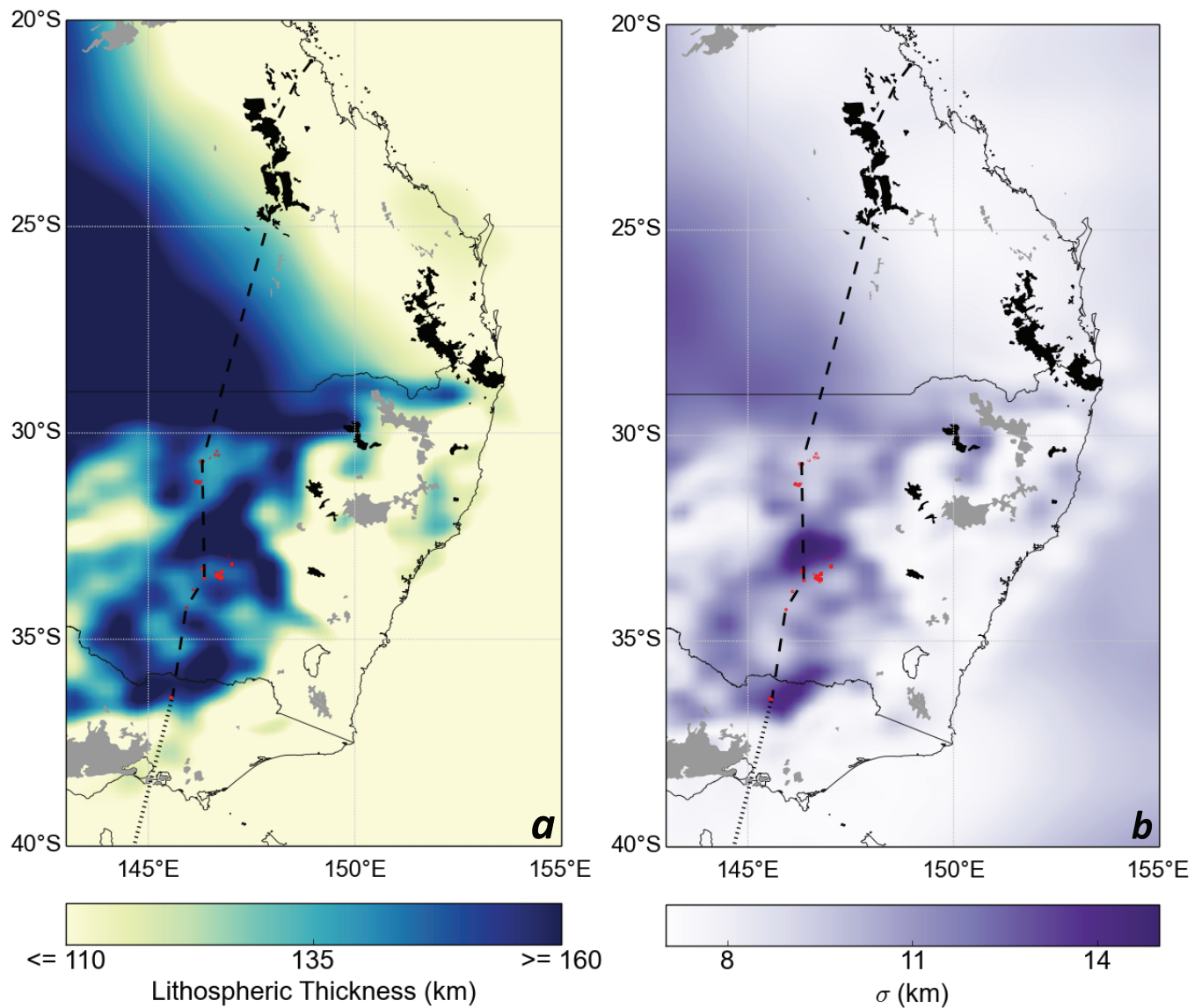
Note that the reconstructions illustrated in Fig. 1a and Extended Data Fig. 2 assume a plume drift velocity of 1 cm yr^{-1} and a melt region diameter of 100 km (black square).



Extended Data Figure 4 | Location of WOMBAT array stations used to create the three-dimensional P-wave velocity model from which our lithospheric thickness estimate was derived. Station spacing is ~ 50 km, which roughly equates to the maximum horizontal resolution of the three-dimensional velocity model.



Extended Data Figure 5 | Depth slice at 120 km, through the three-dimensional P-wave velocity model. North of $\sim 28^\circ$ S, the model reverts to the AuSREM mantle model, owing to a lack of additional data coverage in this region (see Extended Data Fig. 4).



Extended Data Figure 6 | Lithospheric thickness estimate and associated uncertainty. **a, b,** Lithosphere thickness model illustrated in Fig. 1b (**a**), alongside an estimate of its uncertainty (**b**), given by the standard deviation (σ) of an ensemble of 540 plausible models examined. Note that south of

~28°S, the lithospheric thickness estimate is constrained by high-resolution body-wave tomography (~50 km horizontal resolution), whereas north of this latitude it is constrained entirely by the AuSREM mantle model (~200–250 km horizontal resolution)²⁴.

Extended Data Table 1 | Age estimates, derived via ^{40}Ar – ^{39}Ar geochronology, for the volcanic centres considered in this study

Location	Sample ID	Latitude (°)	Longitude (°)	Age (Ma)	Source
Mt Jukes, Hillsborough, QLD	BC-98	-21.00	148.94	33.6 ± 0.5	Ref. 6
Mt Britton, Nebo, QLD	DR13704	-21.47	148.59	32.6 ± 0.4	Ref. 6
Mt Pollux, N. Peak Range, QLD	BC-88	-22.48	147.87	30.7 ± 0.4	Ref. 6
Lords Table Mountain, N. Peak Range, QLD	BC-90	-22.66	148.02	30.4 ± 0.5	Ref. 6
Wolfgang Peak, N. Peak Range, QLD	BC-89	-22.55	147.83	30.3 ± 0.4	Ref. 6
Malvern Hill, S. Peak Range, QLD	BC-94	-22.88	148.22	29.0 ± 0.4	Ref. 6
Ropers Peak, S. Peak Range, QLD	BC-95r	-22.87	148.22	28.6 ± 0.4	Ref. 6
Mt St. Peter, Springsure, QLD	BC-71	-24.00	148.04	28.1 ± 0.3	Ref. 6
Hervey's Knob, Buckland, QLD	BC-57	-24.84	147.92	27.3 ± 0.4	Ref. 6
El Capitan, NSW	CPT II-I	-31.22	146.20	17.9 ± 0.3	Ref. 5
Byrock Quarry, NSW	BQ-7	-30.71	146.31	17.1 ± 0.3	Ref. 5
Begargo Hill, NSW	BEH-V	-33.53	146.36	15.5 ± 0.5	Ref. 5
Flagstaff Hill, NSW	FGH-I	-33.80	146.09	15.3 ± 0.2	Ref. 5
Griffith Quarry, NSW	GR-7	-34.23	145.92	14.9 ± 0.4	Ref. 5
Cosgrove Quarry, NSW	E7980	-36.34	145.60	8.9 ± 0.2	Ref. 5

For illustrative purposes, only nine volcanic centres are shown in Fig. 1a (those coloured in black here). However, all 15 centres were used in our reconstruction of the Cosgrove hotspot track (including those shown in grey here).

Extended Data Table 2 | Rock type, sample locations, sample numbers and data source, from previously analysed samples along the Cosgrove track

Rock-type	Sample location	Sample ID	Source
Leucitite	Weeber Hill, Condobolin, NSW	CND4	Ref. 17
Leucitite	Cosgrove, VIC	CG1	Ref. 26
Leucitite	Cosgrove, VIC	CG2	Ref. 26
Leucitite	Cosgrove, VIC	CG3	Ref. 26
Transitional Basalt	Springsure, QLD	Q135	Ref. 17
Ol-Tholeiite	Springsure, QLD	Q137	Ref. 17
Basanite	Mt. Llandillo, Anakie, QLD	AK16/6	Ref. 17
Basanite	Black Peak, Anakie, QLD	AK39/1	Ref. 17
Leucitite	Begargo-Hill, NSW	GA-3471	Ref. 42
Leucitite	Flagstaff, NSW	GA-3481	Ref. 42
Leucitite	El Capitan, NSW	GA-3479	Ref. 42

Associated trace-element concentrations are provided in Extended Data Table 3. The limited trace-element concentrations provided for the three samples from ref. 42 (shaded grey) are not plotted in Fig. 2.

Extended Data Table 3 | Trace-element concentrations for a number of samples along the Cosgrove track

Element	PM	CND4	CG1	CG2	CG3	Q135	Q137	AK16/6	AK39/1	GA-3471	GA-3481	GA-3479
Cs	0.032	1.05	1.14	1.69	1.05	0.35	0.40	0.20	0.20	-	-	-
Rb	0.635	96.0	90.9	112.1	86.7	23.3	21.0	28.5	26.8	78.9	185	265
Ba	6.989	1150.	966.0	1011.	964.	335.	267.	373.0	454.0	-	-	-
Th	0.095	8.1	8.87	9.26	8.63	4.3	2.9	4.3	7.2	-	-	-
U	0.021	1.3	1.72	1.9	2.09	0.5	0.7	0.7	1.8	-	-	-
Nb	0.713	96.0	107.	118.	116.	41.0	27.0	54.0	60.0	-	-	-
Ta	0.041	6.2	5.44	5.78	5.83	1.8	1.3	2.8	2.5	-	-	-
La	0.687	75.0	79.5	84.6	85.2	26.5	22.5	37.5	50.0	-	-	-
Ce	1.775	158.0	162.	168.	169.	57.0	48.0	76.0	89.0	-	-	-
Sr	21.10	1154.	1131.	1209.	1089.	569.	432.	777.	923.0	1251	1265	972
Nd	1.354	74.0	75.6	80.6	80.6	28.0	24.0	35.5	35.5	86.6	92.7	118.3
Sm	0.444	12.8	13.1	14.3	14.3	5.3	4.0	6.6	7.3	14.1	15.1	16.7
Zr	11.20	536.	433.	470.	463.	233.	180.	219.	192.	-	-	-
Hf	0.309	12.0	9.19	9.8	9.82	4.2	3.6	4.0	3.3	-	-	-
Eu	0.168	3.75	3.88	4.20	4.18	2.05	1.75	2.45	2.3	-	-	-
Gd	0.596	9.60	11.5	12.0	12.2	5.4	4.8	6.6	6.0	-	-	-
Tb	0.108	1.20	1.37	1.46	1.44	0.80	0.61	1.0	1.0	-	-	-
Y	4.550	26.0	31.4	33.5	31.8	24.8	24.0	23.8	21.5	-	-	-
Ho	0.164	1.05	1.08	1.16	1.13	0.70	0.86	1.15	1.0	-	-	-
Yb	0.493	1.55	1.75	1.86	1.78	1.70	1.80	1.90	1.7	-	-	-
Lu	0.074	0.20	0.239	0.248	0.248	0.25	0.27	0.26	0.24	-	-	-

See Extended Data Table 2. Concentrations are in p.p.m. Although not plotted in Fig. 2, the limited trace-element concentrations (Rb, Sr, Nd and Sm) for leucitites provided in ref. 42 (that is, GA-3471, GA-3479 and GA-3481) are consistent with the concentrations provided in the more complete data sets of refs 17 and 26. PM denotes primitive mantle concentrations from ref. 31.

Novel competitors shape species' responses to climate change

Jake M. Alexander¹, Jeffrey M. Diez² & Jonathan M. Levine¹

Understanding how species respond to climate change is critical for forecasting the future dynamics and distribution of pests, diseases and biological diversity^{1–3}. Although ecologists have long acknowledged species' direct physiological and demographic responses to climate, more recent work suggests that these direct responses can be overwhelmed by indirect effects mediated via other interacting community members^{2–7}. Theory suggests that some of the most dramatic impacts of community change will probably arise through the assembly of novel species combinations after asynchronous migrations with climate^{8–10}. Empirical tests of this prediction are rare, as existing work focuses on the effects of changing interactions between competitors that co-occur today^{7,11–15}. To explore how species' responses to climate warming depend on how their competitors migrate to track climate, we transplanted alpine plant species and intact plant communities along a climate gradient in the Swiss Alps. Here we show that when alpine plants were transplanted to warmer climates to simulate a migration failure, their performance was strongly reduced by novel competitors that could migrate upwards from lower elevation; these effects generally exceeded the impact of warming on competition with current competitors. In contrast, when we grew the focal plants under their current climate to simulate climate tracking, a shift in the competitive environment to novel high-elevation competitors had little to no effect. This asymmetry in the importance of changing competitor identity at the leading versus trailing range edges is best explained by the degree of functional similarity between current and novel competitors. We conclude that accounting for novel competitive interactions may be essential to predict species' responses to climate change accurately.

Climate change will alter species' competitive environments through initial shifts in the performance and relative abundance of their current competitors, and longer-term changes in the identity of their competitors caused by migration and local extinctions². Empirical studies of the shorter-term changes in neighbour abundance provide evidence both for^{7,11,13,15} and against^{12–14} the importance of competitive interactions in mediating the impact of climate change. However, these results may underestimate the potential role of changing competition. Over longer timescales, species will experience competition from new and functionally different migrants, and if they themselves migrate to track climate change, they will probably encounter new resident competitors^{2,9}. Despite the potential importance of these novel competitive interactions in determining species' persistence and future distributions with climate change^{16,17}, empirical evidence is scant for two reasons. First, in most systems, the combinations of species that will face one another in the future is highly uncertain. Second, the logistical challenges associated with experimentally assembling hypothetical future communities, and doing so under realistic climate scenarios, are typically prohibitive.

Elevation gradients in mountains provide a unique opportunity to test how changing competitor identity will affect species' responses to climate change. The steep climate gradient in these environments

means that the novel competitors that species will face following climate warming are those already occurring only hundreds of metres away. Furthermore, perennial grasslands in these regions lend themselves to whole-community transplantation along climate gradients. We experimentally simulated the endpoints of the spectrum of competitive environments that an alpine species will experience following climate change at the leading and trailing edges of its range (Fig. 1). At its trailing range edge, a species that fails to migrate will experience warmer climate and compete with either its current community members (scenario 1 in Fig. 1), or with a novel community composed of species that have migrated upwards from lower elevation (scenario 2). By contrast, at the leading edge of its range, a species migrating to higher elevations to track its current climate will compete either with its current competitors if they also migrate (scenario 3) or with a novel higher-elevation community that has persisted in place (scenario 4).

To simulate these scenarios, we transplanted focal alpine species and intact plant communities along an elevation gradient in the Swiss Alps (Table 1), and followed their performance for 2 years. To simulate scenarios in which focal species and/or communities fail to migrate and thus experience warmer temperatures, we moved focal plants and/or communities to a lower-elevation site. To simulate scenarios in which focal species and/or communities migrate to track current climate and thus experience little change in temperature, we transplanted them back into their current elevation site. The direction of transplantation is thus meant to reflect future climate conditions, not

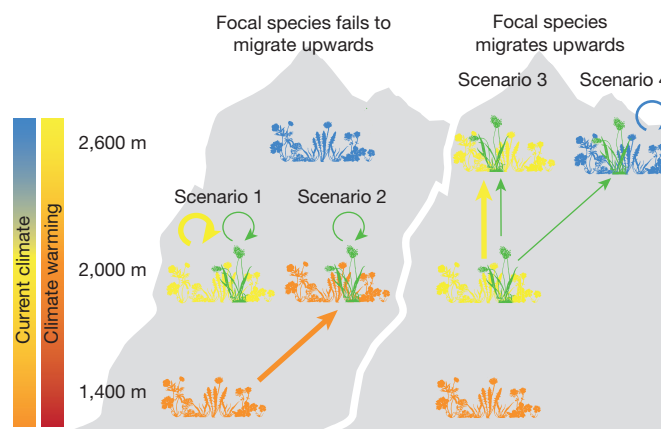


Figure 1 | Scenarios for the competition experienced by a focal alpine plant following climate warming. If the focal plant species (green) fails to migrate, it competes either with its current community (yellow) that also fails to migrate (scenario 1) or, at the other extreme, with a novel community (orange) that has migrated upwards from lower elevation (scenario 2). If the focal species migrates upwards to track climate, it competes either with its current community that has also migrated (scenario 3) or, at the other extreme, with a novel community (blue) that has persisted (scenario 4). Table 1 describes the experimental implementation of these scenarios.

¹Institute of Integrative Biology, ETH Zurich, Universitätsstrasse 16, 8092 Zurich, Switzerland. ²Department of Botany and Plant Sciences, University of California Riverside, 900 University Avenue, Riverside, California 92521, USA.

Table 1 | Experimental manipulations corresponding to the different competitive scenarios experienced by a focal plant following climate warming

Scenario	Focal species' response to warming	Origin of focal species	Competitor scenario	Origin of competitors	Elevation of transplant site
1	Focal species fails to migrate and experiences warming	2,000 m	Current competitors persist in warmer climate	2,000 m	1,400 m
2	Focal species fails to migrate and experiences warming	2,000 m	Low-elevation competitors migrate up and replace current competitors	1,400 m	1,400 m
3	Focal species migrates up to track climate	2,000 m	Current competitors migrate up to track climate and replace high-elevation competitors	2,000 m	2,000 m
4	Focal species migrates up to track climate	2,000 m	High-elevation competitors persist in warmer climate	2,600 m	2,000 m

the future location of the species (Table 1). Plants moved downhill experienced an average daily climate warming of around 3 °C (Extended Data Fig. 1 and Extended Data Table 1), which reflects the magnitude of climate change predicted for the next 50–100 years in Switzerland¹⁸. While abrupt climate change experiments, as imposed here, mimic future conditions¹⁹, testing more gradual species' responses, such as adaptation, requires other approaches. We tested the influence of the four migration scenarios on the performance of four focal alpine species: *Anthyllis vulneraria* ssp. *alpestris* (alpine kidney vetch, hereafter *A. alpestris*), *Plantago atrata* (black plantain), *Pulsatilla vernalis* (spring

pasqueflower) and *Scabiosa lucida* (glossy scabious). These species differ in their dispersal potential (Extended Data Table 2), and their current ranges do not effectively extend to either the lowest- or highest-elevation field sites (see Methods and Table 1).

The response of the focal species to novel competitors depended on whether they grew at the experimental site with warmer or current climate conditions (Fig. 2; significant novel competitor × site, or novel competitor × site × species interactions in Table 2). When the focal species experienced increased temperature (transplantation to lower elevation to simulate climate warming at the trailing edge of their

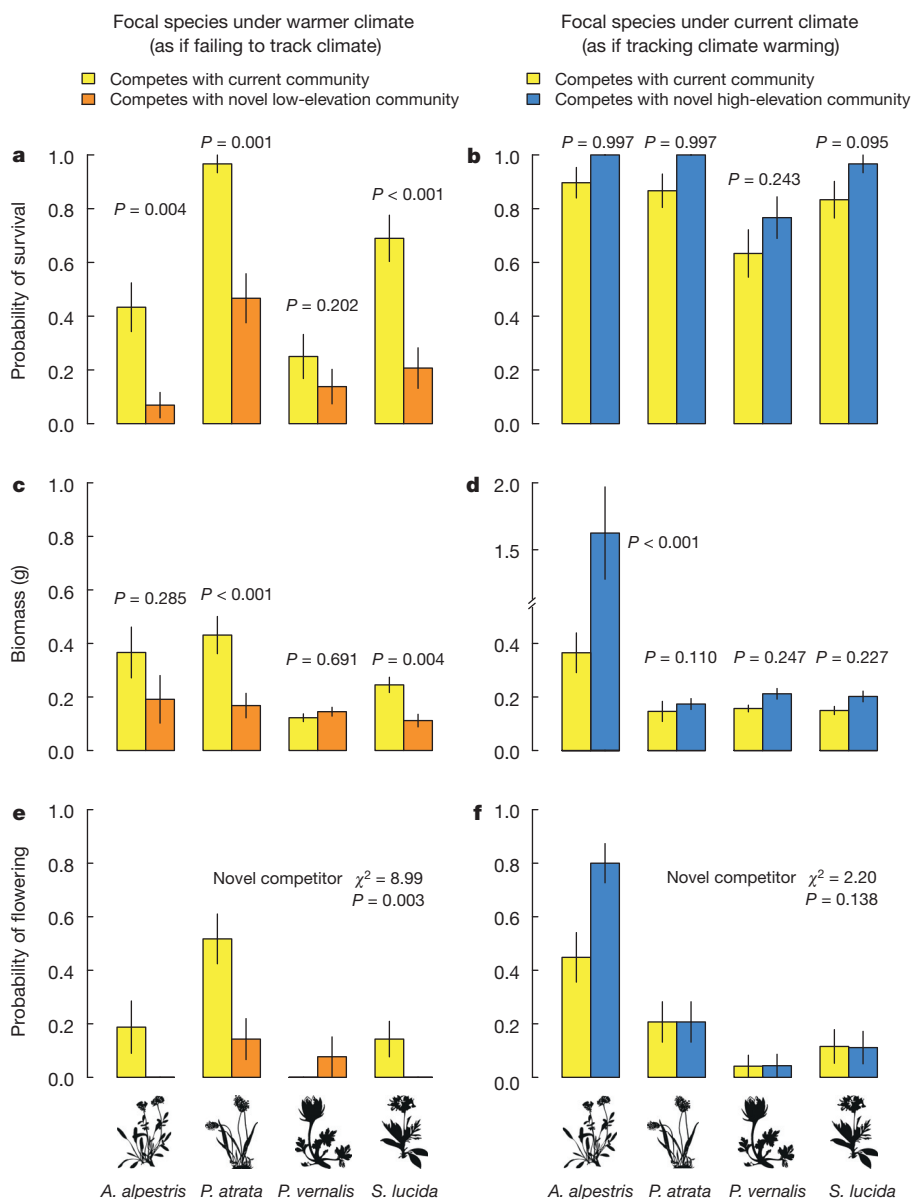


Figure 2 | Effect of novel competitors on alpine plant performance. Survival over 2 years (a, b), second year biomass (c, d) and second year flowering (e, f) of focal species exposed to different potential competition scenarios following climate warming (see Table 1). Shown are means (s.e.m.) of the raw data. When the novel competitor × site × species interaction was significant (a–d), *P* values for species-by-site specific contrasts were taken from the full model (see Table 2 for statistics and *n*), else from site-specific contrasts averaging over species (e, f); *P* values < 0.005 remain significant ($\alpha = 0.05$) after Holm–Bonferroni correction for multiple comparisons.

Table 2 | Statistical analysis of focal alpine plant performance

Source	d.f.	Survival		Biomass		Flowering	
		χ^2	P	χ^2	P	χ^2	P
Novel competitor (NC)	1	2.53	0.112	0.01	0.944	0.03	0.861
Site (S)	1	32.38	<0.001	<0.01	0.991	1.08	0.298
Species (Sp)	3	67.21	<0.001	55.98	<0.001	61.87	<0.001
NC \times S	1	28.55	<0.001	26.04	<0.001	5.28	0.022
NC \times Sp	3	0.93	0.818	14.08	0.003	5.64	0.131
S \times Sp	3	12.34	0.006	18.81	<0.001	17.41	0.001
NC \times S \times Sp	3	10.14	0.017	10.16	0.017	5.02	0.170
n (n blocks)		473 (20)		291 (20)		363 (20)	

Shown are likelihood ratio tests for novel competitor (current versus novel competitors), site (1,400 or 2,000 m experimental site), and species effects and their interactions on survival after 2 years, and biomass and flowering probability in the second year of the experiment. Also shown are the total number of observations (n) and experimental units (n blocks) for each model.

range), their performance 2 years after transplantation depended strongly on the origin of their competitors (Fig. 2). For three of four species, survival was reduced by 52–84% (Fig. 2a), biomass by 48–61% (Fig. 2c; n.s. for *A. alpestris*) and flowering by over 72% (Fig. 2e) when competing against a novel, low-elevation plant community (scenario 2) compared with their current alpine community (scenario 1). The biomass reduction due to these potential migrants from lower elevation was significant even in the first year of the study (Extended Data Fig. 2; novel competitor $\chi^2 = 17.66$, d.f. = 1, $P < 0.001$). We found much weaker effects of changing competitor identity when focal species were transplanted back into their current elevation to simulate migration and climate tracking. Here, whether focal species competed with a novel high-alpine community (scenario 4) or their current community (scenario 3) had no significant effect on survival (Fig. 2b) or flowering (Fig. 2f), and modest, largely non-significant effects on biomass (Fig. 2d). The one exception was the strong response of *A. alpestris* biomass to novel competitors, but this response was replicated when it grew without any competitors on the soils from the two elevations ($\chi^2 = 7.31$, d.f. = 1, $P = 0.007$; Extended Data Fig. 3), suggesting a limited role for shifting competitor identity.

Plant performance in our experiment might be affected by factors other than competitor identity that differ between the communities, including soil chemistry and biota. To evaluate soil effects on plant growth, we grew the focal species at each site without competition on soil originating from each elevation. We found that focal species tended to grow better at lower- versus higher-elevation (site $\chi^2 = 24.31$, d.f. = 1, $P < 0.001$) on a common 2,000 m soil, but their response to soil origin never matched significant biomass responses to novel competitors, suggesting that the observed changes in performance in Fig. 2 were indeed due to shifting plant competition (with the exception of *A. alpestris* as mentioned above; Extended Data Fig. 3). We also conducted a follow-up greenhouse experiment to isolate the effects of soil biota from different elevations. Soil organisms could affect plant competition if they fail to migrate synchronously with the plant communities in the future. Results suggest that soil biota from the different elevations did not affect the relative performance of alpine versus sub-alpine plant competitors (Extended Data Fig. 4 and Extended Data Table 3). Related to this, we did not find differences in the incidence of herbivory across the two community types at the low-elevation site (where competitor identity effects were strong), except for two species in the first year of the study only (and this did not relate to subsequent survival or biomass; Extended Data Table 4).

In sum, our results show that novel competitors strongly affected the performance of alpine plants under increased temperatures, as will occur at the trailing edge of their range, but had little effect on plants under current temperatures, as would occur following range expansion to higher elevation. This asymmetry in the importance of competitor identity at the leading versus trailing range edges can be explained by the greater functional similarity between the high- and middle-elevation communities, measured with field-based trait measurements on 61 species. The low- and middle-elevation communities were

2.4 times further apart along the first principal component of trait space (Fig. 3a; explaining 76% of the variation in community-weighted trait means) and their comparison produced seven times the *F* statistic in a permutation-based multivariate analysis of variance (MANOVA) ($F_{1,27} = 52.94$ versus 7.56, $P < 0.001$ and $P < 0.01$, respectively) than the middle- and high-elevation communities. The greater functional similarity of the two higher-elevation communities was caused by shared functional traits, particularly leaf size, leaf mass and plant height, not shared species; the communities were equally distinct in their species composition (Fig. 3b; 1,400 versus 2,000 m permutation MANOVA, $F_{1,27} = 27.74$; 2,000 versus 2,600 m $F_{1,27} = 24.63$; $P < 0.001$ for both comparisons).

Finally, first-year biomass results (before heavy mortality in the second year of the experiment) allowed us to compare the effect of warming on our focal species' interactions with current competitors with the effect of community changes that will arise from competitor migration and local extinction. We found that when the focal alpine plants grew with their current alpine community under warmer temperatures, they experienced greater competition than under current temperatures (pink versus white bars in Fig. 4), but these effects were weak and not significant for any species. This result is consistent with the mixed results from previous studies of short-term competitor dynamics under climate change^{7,11–15}. By contrast, for *P. atrata* and *S. lucida*, significantly greater effects of competition arose from changing competitor identity and warmer conditions (Fig. 4). These are also two of the three species with traits predicting relatively poor dispersal (Extended Data Table 2) and thus interactions with novel low-elevation competitors may determine their eventual persistence. This result further suggests that the strongest effects of climate change on competition in this system are likely to occur after the immigration of novel competitors at species' trailing range edges.

Our study provides some of the first empirical evidence that accounting for novel competitors may be important to predicting species' responses to climate change^{17,20}. Specifically, our results suggest that species' range dynamics probably depend not only on their ability to track climate, but also the migration of their competitors, and the extent to which novel and current competitors exert differing competitive effects. In our system, populations might persist at their trailing range edge in areas soon to be warmer, as long as lower-elevation migrants fail to arrive. This prediction parallels results from the few mechanistic studies of population decline following climate change, where changing biotic interactions appear more important than direct physiological effects of warmer temperature¹. However, our results also suggest that, in some cases, changing competitor identity may be less important. We found, for example, that the shift to

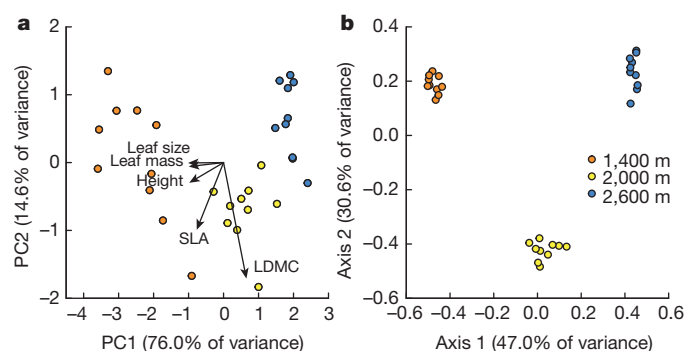


Figure 3 | Functional and floristic community composition. Ordinations of ten replicate communities from sites at 1,400 m, 2,000 m and 2,600 m elevation based on (a) a principal component (PC) analysis of community-weighted means of five functional traits (SLA, specific leaf area; LDMC, leaf dry matter content), and (b) a principal coordinates analysis of floristic composition based on Bray–Curtis dissimilarity. In a, arrows show the loading of individual traits on each principal component axis (loadings have been multiplied by 2 for clarity).

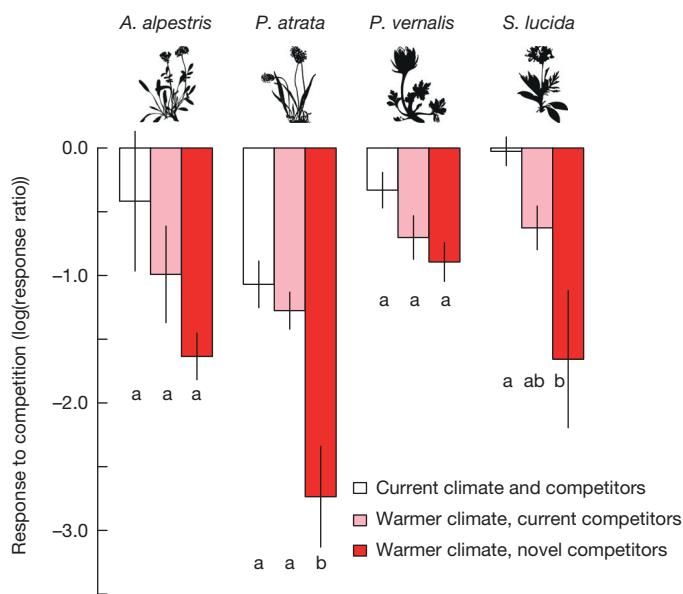


Figure 4 | The response of four alpine species to competition. Plants grew with either (1) their current competitors and climate (in a site at 2,000 m), (2) their current competitors and warmer climate (growing at 1,400 m) or (3) novel competitors from low elevation and warmer climate (growing at 1,400 m). Shown are mean log response ratios (s.e.m.) of above-ground biomass calculated from plants growing with or without competitors ($n = 25, 30, 27, 29$, for each species, respectively). Different letters below the bars for each species indicate significantly different contrasts (Tukey's honest significant difference test, $P < 0.05$).

novel high-alpine competitors is unlikely to influence the range expansion of focal species to higher elevation, in agreement with the rapid migration of many species upslope with recent climate warming^{21,22}. Future work combining species' functional traits, detailed distribution information and ecological theory may prove particularly useful for forecasting how novel competitive interactions determine the response of biological diversity to climate change^{23,24}.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 2 May 2014; accepted 16 July 2015.

Published online 16 September 2015.

1. Cahill, A. E. *et al.* How does climate change cause extinction? *Proc. R. Soc. B* **280**, 20121890 (2012).
2. Gilman, S. E., Urban, M. C., Tewksbury, J., Gilchrist, G. W. & Holt, R. D. A framework for community interactions under climate change. *Trends Ecol. Evol.* **25**, 325–331 (2010).
3. Tylianakis, J. M., Didham, R. K., Bascompte, J. & Wardle, D. A. Global change and species interactions in terrestrial ecosystems. *Ecol. Lett.* **11**, 1351–1363 (2008).
4. Davis, A. J., Jenkinson, L. S., Lawton, J. H., Shorrocks, B. & Wood, S. Making mistakes when predicting shifts in species range in response to global warming. *Nature* **391**, 783–786 (1998).

5. González-Megías, A. & Menéndez, R. Climate change effects on above- and below-ground interactions in a dryland ecosystem. *Phil. Trans. R. Soc. B* **367**, 3115–3124 (2012).
6. Liancourt, P. *et al.* Plant response to climate change varies with topography, interactions with neighbors, and ecotype. *Ecology* **94**, 444–453 (2013).
7. Suttle, K. B., Thomsen, M. A. & Power, M. E. Species interactions reverse grassland responses to changing climate. *Science* **315**, 640–642 (2007).
8. Blois, J. L., Zarnetske, P. L., Fitzpatrick, M. C. & Finnegan, S. Climate change and the past, present, and future of biotic interactions. *Science* **341**, 499–504 (2013).
9. Urban, M. C., Tewksbury, J. J. & Sheldon, K. S. On a collision course: competition and dispersal differences create no-analogue communities and cause extinctions during climate change. *Proc. R. Soc. B* **279**, 2072–2080 (2012).
10. Williams, J. W. *et al.* Model systems for a no-analog future: species associations and climates during the last deglaciation. *Ann. NY Acad. Sci.* **1297**, 29–43 (2013).
11. Adler, P. B., Dalgleish, H. J. & Ellner, S. P. Forecasting plant community impacts of climate variability and change: when do competitive interactions matter? *J. Ecol.* **100**, 478–487 (2012).
12. Adler, P. B., Leiker, J. & Levine, J. M. Direct and indirect effects of climate change on a prairie plant community. *PLoS One* **4**, e6887 (2009).
13. Farrer, E. C., Ashton, I. W., Knape, J. & Suding, K. N. Separating direct and indirect effects of global change: a population dynamic modeling approach using readily available field data. *Glob. Change Biol.* **20**, 1238–1250 (2014).
14. Levine, J. M., McEachern, A. K. & Cowan, C. Do competitors modulate rare plant response to precipitation change? *Ecology* **91**, 130–140 (2010).
15. Milazzo, M., Mirto, S., Domenici, P. & Gristina, M. Climate change exacerbates interspecific interactions in sympatric coastal fishes. *J. Anim. Ecol.* **82**, 468–477 (2013).
16. HilleRisLambers, J., Harsch, M. A., Ettinger, A. K., Ford, K. R. & Theobald, E. J. How will biotic interactions influence climate change-induced range shifts? *Ann. NY Acad. Sci.* **1297**, 112–125 (2013).
17. Wisz, M. S. *et al.* The role of biotic interactions in shaping distributions and realised assemblages of species: implications for species distribution modelling. *Biol. Rev. Camb. Philos. Soc.* **88**, 15–30 (2013).
18. CH2011. *Swiss Climate Change Scenarios CH2011* (C2SM, MeteoSwiss, ETH Zurich, NCCR Climate and OCC, 2011).
19. Elmendorf, S. C. *et al.* Experiment, monitoring, and gradient methods used to infer climate change effects on plant communities yield consistent patterns. *Proc. Natl Acad. Sci. USA* **112**, 448–452 (2015).
20. Svenning, J.-C. *et al.* The influence of interspecific interactions on species range expansion rates. *Ecography* **37**, 1198–1209 (2014).
21. Chen, I.-C., Hill, J. K., Ohlemüller, R., Roy, D. B. & Thomas, C. D. Rapid range shifts of species associated with high levels of climate warming. *Science* **333**, 1024–1026 (2011).
22. Gottfried, M. *et al.* Continent-wide response of mountain vegetation to climate change. *Nature Clim. Change* **2**, 111–115 (2012).
23. Adler, P. B., Fajardo, A., Kleinhesselink, A. R. & Kraft, N. J. B. Trait-based tests of coexistence mechanisms. *Ecol. Lett.* **16**, 1294–1306 (2013).
24. Freckleton, R. P. & Watkinson, A. R. Predicting competition coefficients for plant mixtures: reciprocity, transitivity and correlations with life-history traits. *Ecol. Lett.* **4**, 348–357 (2001).

Acknowledgements We thank M.-J. Mächler, D. Righetti, C. Schmid, P. Stettler, R. Guidon, A. Vitra, S. Minneboo, J. Leuenberger and other members of the Plant Ecology group for assistance with field work, and the community of Haldenstein for providing field sites. S. Güsewell provided statistical advice. We thank P. Adler, J. HilleRisLambers and the Plant Ecology group for reading and commenting on the manuscript. ETH Zurich funding to the Plant Ecology group supported the project.

Author Contributions All authors designed the study, assisted with fieldwork and wrote the paper. J.M.A. analysed the data and wrote the first draft.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to J.M.A. (jake.alexander@usys.ethz.ch).

METHODS

Field transplant experiment. We selected three perennial grassland sites (1,400, 2,000 and 2,600 m above sea level) in the Swiss Alps (Calanda mountain, Canton des Grisons), which are all dominated by compact turfs, and contain different, overlapping sets of species (Fig. 3). The sites are at maximum 3 km apart with similar southeast exposure, slope and calcareous bedrock, but span a steep climate gradient, with a temperature range of 6 °C from subalpine (1,400 m) to higher alpine (2,600 m) sites, as measured over the duration of the experiment (Extended Data Fig. 1, Extended Data Table 1). The temperature differences between the lower and middle sites, and the upper and middle sites, during the experiment were on average 2.6 °C and 3.4 °C, respectively. Precipitation changes were smaller but decreased by approximately 16% with both 600 m drops in elevation (Extended Data Table 1), also consistent with expectations for climate change¹⁸. The lower and middle sites are managed as summer pasture, and the upper site is grazed by native ungulates.

At the end of August 2012, 75 cm × 75 cm turfs containing intact plant communities, including roots and the organic soil layer, were excavated at each site to a depth of 20 cm. To implement the design in Table 1, the site at 1,400 m received ten transplanted communities from the 2,000 m site and ten communities transplanted from other locations at the 1,400 m site. Meanwhile the site at 2,000 m received ten transplanted communities from the 2,600 m site and ten from other locations at the 2,000 m site (Table 1). Soil was obtained from each site (after removing the vegetation) and transplanted across sites in the same design. At the two transplant destination sites, each treatment (two communities and two soils) was assigned at random to one of four plots within each of ten blocks (giving 80 plots in total). Blocks were separated by 1 m, with 0.5 m between treatments within blocks.

Focal individuals of four alpine species (*A. vulneraria* ssp. *alpestris*, *P. atrata*, *P. vernalis* and *S. lucida*) were obtained by cutting 240 plugs (about 3 cm diameter) containing a single adult plant from the 2,000 m site. These species are widespread at the 2,000 m site and are either not found or extremely rare in the communities at the 1,400 m and 2,600 m sites. Three individuals per species were planted at random, 15 cm apart in a grid within each treatment and block ($n = 30$ per species, treatment and site). The lower two transplant sites were fenced to exclude cattle, as well as marmots at the 2,000 m site (marmots are not seen at the other sites).

To minimize transplantation-related issues, we transplanted the communities and focal plants in late summer/early autumn, after the plants had already begun to senesce, so that they would first experience their new climate in their growth phase when they emerged the following spring. The communities were clipped to reduce evapo-transpirative stress during transplantation, and the transplants were watered and protected with shade cloth for 1 week after transplantation. Any focal plants that died within the first 3 weeks were replaced. We note that focal individuals and communities did not respond to transplantation in ways that would suggest they were poorly adapted to climate or other conditions at lower elevation: 2,000 m focal plants growing without competition on soils from 2,000 m were larger when transplanted downslope than when transplanted to the same elevation, and grew better on the soil from the lower elevation (1,400 m) site (Extended Data Fig. 3). The above-ground biomass of the intact transplanted communities was unaffected by transplantation downslope (Extended Data Fig. 5).

The survival, phenology and number of inflorescences of every focal individual were monitored every 2 weeks after snowmelt in 2013 and 2014 (deaths occurring in 2014 were confirmed by a final check for surviving plants in early summer 2015). We do not report flowering incidence in the first year of the study because flowering in many alpine plants is determined by conditions in the preceding year²⁵ (results were generally non-significant). Each focal individual's leaf number and longest leaf length were recorded at planting and in August/September 2013; leaf number and the average area of the three largest leaves were measured in September 2014. Above-ground biomass of each focal individual was estimated from linear models predicting the biomass of destructively harvested individuals from outside the experiment ($n = 27$ – 39), dependent on the non-destructive measures in either 2013 or 2014, including maximum number of inflorescences in 2014 for all species except *P. vernalis* ($R^2 = 0.80$ – 0.98). Regressions were forced through the origin, and negative biomass estimates (possible with interactions) set to 0.01 g. Individuals that died up to a month after snowmelt at each site in spring 2013 were considered to have died of transplant shock and excluded from further analysis. Community above-ground biomass was estimated towards the end of each summer using pin quadrats calibrated with destructive harvests made at each site ($n = 20$ plots at the 1,400 m site, $n = 18$ at 2,000 m, $n = 10$ at 2,600 m). The composition and cover of species in each replicate community were determined in 2013 in mid-May (1,400 m) and mid-June (2,000 m), and again in late August/early September (all sites). Temperature and light intensity were recorded at 30-min intervals using at least one HOBO Pendant data logger (UA-002-64, onset,

www.onsetcomp.com/) at each site. At the end of each season the communities were clipped to approximate biomass removal by grazers.

Statistical analysis of field experiment. No statistical methods were used to predetermine sample size. The investigators were not blinded to allocation during experiments and outcome assessment.

The effects of novel competitors on the biomass of focal plants transplanted into competitor communities from different locations were analysed with mixed-effects models fitted by maximum likelihood. The full model contained main effects and all interactions of the 'novel competitor' treatment (novel versus current competitors), 'site' (1,400 or 2,000 m experimental sites) and 'species,' as well as initial size, as fixed effects, and plot nested in block as random effects to account for the dependency of observations. Biomass and initial size were log-transformed to meet model assumptions of normality and homogeneity of variances. The ten replicate communities per treatment and site were the biological replicates in our experiment, and this number was chosen to be in excess of previous studies that investigated effects of climate change on competition within communities (for example, refs 7, 14). Survival until the end of 2014 and the probability of flowering in 2014 were analysed in a similar way, but using generalized linear mixed models with a binomial family. We tested for over-dispersion (that is, clustering of the binary outcomes within species within plots) and found no significant contribution based on a likelihood ratio test. The statistical significance of individual terms was determined by comparing each model with the correspondingly reduced model using likelihood ratio tests. When the novel competitor × site × species interaction was significant, the significance of the novel competitor effect for each species by site combination was obtained from contrasts within the full model. Significant novel competitor × site interactions in the full model (without a significant three-way interaction) were followed by fitting site-specific models, and testing the novel competitor effect using likelihood ratio tests. When the direction of a species' significant biomass response to novel competitors paralleled its response to the bare soil from the two competitor communities, we also tested the effect of soil itself. In such cases, we used likelihood ratio tests to test the effect of soil origin on biomass for the species-site combination of interest. All models were fitted in R²⁶ using the lme4 package.

A different statistical model was used to compare the response of each focal species to competition from its current competitors under current versus warmer climate, or to novel competitors under warmer climate. For each species and block, log response ratios of biomass (that is, $\ln(\text{biomass with competition/biomass without competition})$) were calculated on the basis of the biomass of individuals competing with a particular plant community versus growing alone on soil from the same community. Differences in competitive responses between treatments were tested using Tukey's honest significant difference tests within a linear model fitted for each species. We used data from 2013 owing to the high mortality and correspondingly low replication for these tests in 2014.

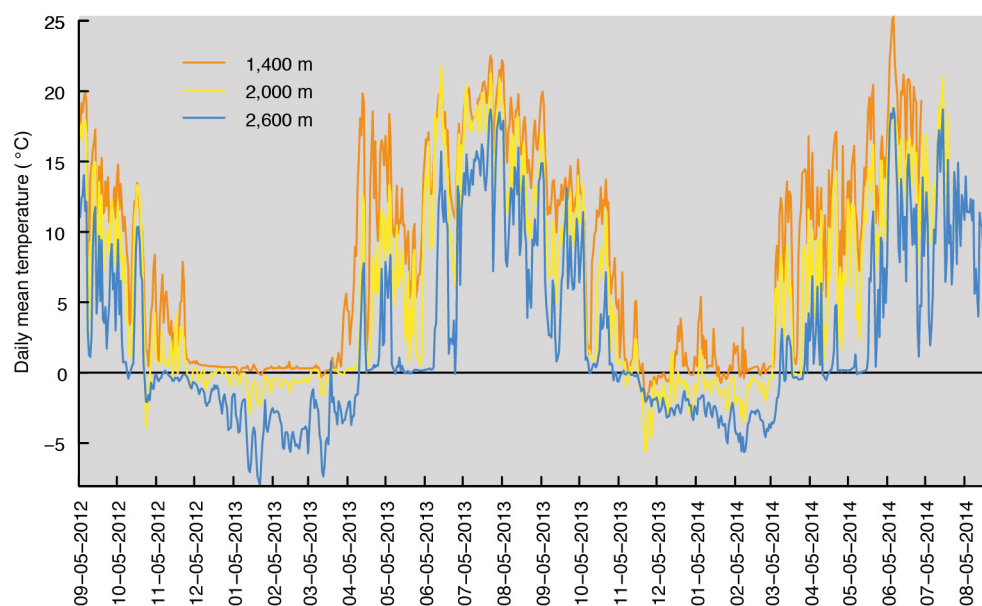
To investigate the functional composition of each community, data on five functional traits (plant vegetative height, specific leaf area ($\text{mm}^2 \text{g}^{-1}$), leaf size (mm^2), leaf mass, leaf dry matter content (mg g^{-1})) were collected from plants growing at the field sites in 2014 using standard methods²⁷ for 61 species that collectively accounted for $89.4 \pm 4.8\%$ (mean \pm s.d.) of the relative cover in these communities ($n = 26$ or 10 per species for height or leaf traits, respectively). Community-weighted means of each trait were calculated by summing the trait values of species within each replicate community, weighted by their relative cover. Differences between pairs of communities (using a subset of ten replicate communities from each elevation to ensure equal sampling effort) in terms of community-weighted functional trait means were analysed with a permutation-based MANOVA (function 'adonis' in the R package 'vegan', using a Euclidean distance matrix), and visualized using a principal components analysis. The same analysis, but based on a Bray–Curtis distance matrix, was applied to community differences in floristic composition in 2013. The effect of site (1,400 m, 2,000 m, 2,600 m) was highly significant in both cases ($F_{2,27} > 26.18$, $P < 0.001$), and these analyses were followed by pairwise contrasts between the high and middle sites and between the middle and low sites as reported in the main text. Compositional differences were visualized using principal coordinates analysis on a Bray–Curtis distance matrix of log-transformed cover values (mean of the two sampling dates) for the same 61 species.

Soil biota experiment. Whether the microbial community at each elevation will migrate in unison with the plants from those elevations is not clear. We therefore conducted a follow-up greenhouse experiment to investigate potential effects of the soil biota from the lower elevation (1,400 m) and alpine (2,000 m) sites on the relative performance of lower versus higher-elevation competitors. We grew three plant species from the 1,400 m site and three 2,000 m focal species (all but *A. alpestris*) with soil inoculum from the 1,400 m and 2,000 m sites. The background soil was a mixture of soil collected at the 1,400 m site and a 2,200 m site,

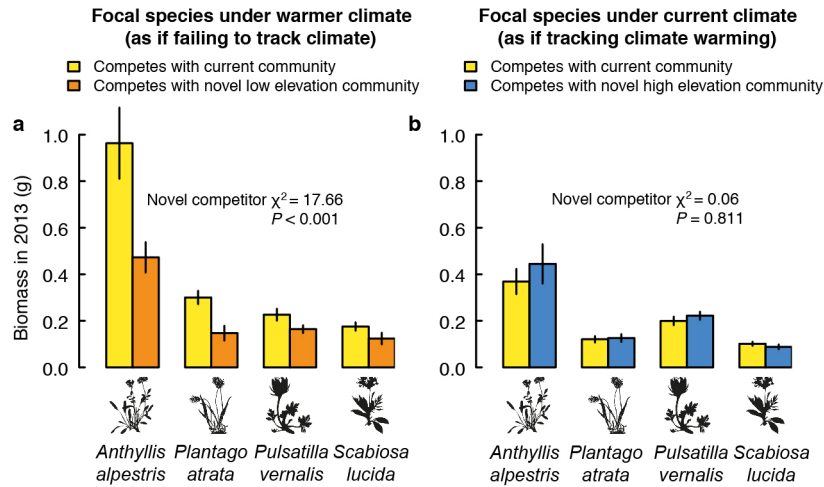
which had been sieved, homogenized and sterilized. We sterilized by autoclaving at 121 °C for 20 min, and again after a 2 day incubation period. Live soil inoculum was collected from soil cores from the top *ca.* 15 cm of the soil profile at the 1,400 and 2,000 m sites in October 2014, and sieved and stored at 4 °C before use. Seedlings were germinated from field-collected seed on filter paper, and then transplanted as a single individual to a 360 ml pot containing sterilized background soil and an inoculum (9% of total soil mass) of live soil from one of the elevations. One pot from each species/soil combination was arranged at random within a block ($n = 10$ blocks) on a single bench in a glasshouse in Zurich, Switzerland (set to 20 °C, 14 h day, with supplementary lighting). Each pot received its own drip tray to minimize cross-contamination during watering. After 3 months, above-ground plant parts were harvested to determine dry mass. For each species, a linear mixed-effects model containing soil community origin as a fixed effect and block as a random effect was fitted by maximum likelihood,

and compared with a simpler model without soil community using a likelihood ratio test.

25. Körner, C. *Alpine Plant Life: Functional Plant Ecology of High Mountain Ecosystems* (Springer, 2003).
26. R Development Core Team. R: a language and environment for statistical computing v.3.1.3 (R foundation for statistical computing, 2015).
27. Pérez-Harguindeguy, N. *et al.* New handbook for standardised measurement of plant functional traits worldwide. *Aust. J. Bot.* **61**, 167–234 (2013).
28. Zimmermann, N. E. & Kienast, F. Predictive mapping of alpine grasslands in Switzerland: Species versus community approach. *J. Veg. Sci.* **10**, 469–482 (1999).
29. Hintze, C. *et al.* D³: the Dispersal and Diaspore database – baseline data and statistics on seed dispersal. *Perspect. Plant Ecol. Evol. Syst.* **15**, 180–192 (2013).
30. Tamme, R. *et al.* Predicting species' maximum dispersal distances from simple plant traits. *Ecology* **95**, 505–513 (2014).

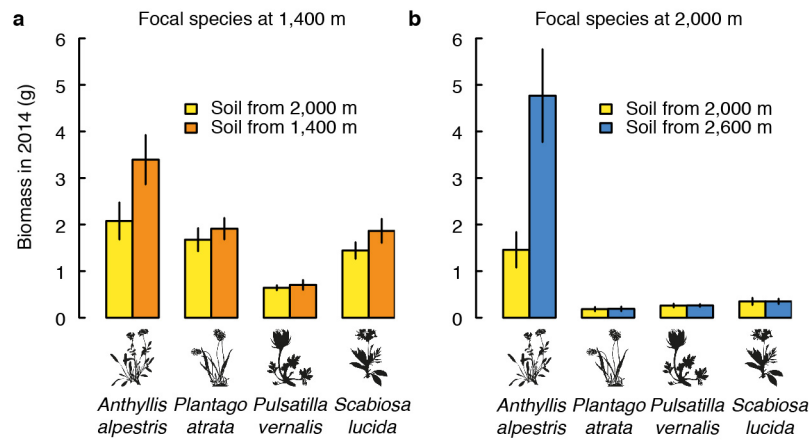


Extended Data Figure 1 | Daily mean temperature during the study at the three experimental sites.



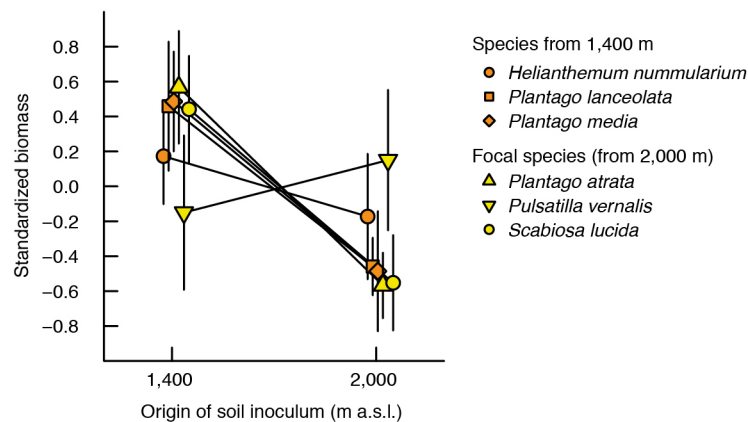
Extended Data Figure 2 | Effect of novel competitors on alpine plant biomass in 2013. Focal species were exposed to different competition scenarios, depending on whether they and/or their surrounding community would either migrate, or fail to migrate, following climate warming (see Fig. 1). Shown are means (s.e.m.) of the raw data, and likelihood ratio tests

(d.f. = 1, $n = 182$ (a) and 221 (b), $n = 10$ experimental units (blocks) per site) of the novel competitor effect at each experimental site (in the main model, across all species and sites: novel competitor \times site interaction $\chi^2 = 8.42$, d.f. = 1, $P = 0.004$; novel competitor \times site \times species interaction $\chi^2 = 3.17$, d.f. = 3, $P = 0.367$).



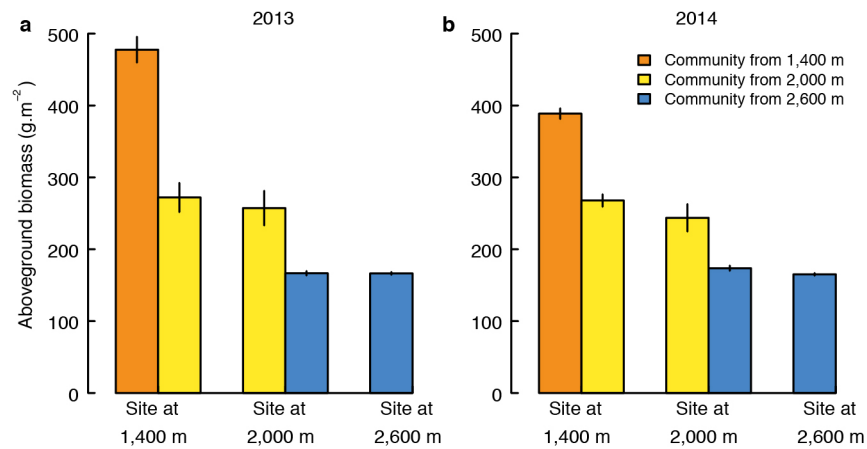
Extended Data Figure 3 | Biomass in 2014 of four alpine plant species growing on soils without competition. Plants grew under a warmer climate (a, at 1,400 m) or under their current climate (b, at 2,000 m), either on soil from

that site, or on soil from a site 600 m higher up the mountain slope. Shown are means (s.e.m.) of the raw data (total $n = 314$).



Extended Data Figure 4 | Effect of soil biota on plant biomass. Plants grew on soils inoculated with soil biota from 1,400 or 2,000 m. Plants grew better with soil biota originating from lower elevation, but this effect was shared across species from 2,000 m (in yellow, focal species from the field experiment)

and 1,400 m (orange). Thus how fast the 1,400 m soil biota migrate or rise to dominance at higher elevation in the future may not strongly determine the relative performance of 1,400 and 2,000 m plants. Shown are means (s.e.m.) of standardized plant biomass. For statistics and n see Extended Data Table 3.



Extended Data Figure 5 | Above-ground community biomass. Standing biomass was estimated in late summer 2013 (a) and 2014 (b) in the plant communities from sites at 1,400, 2,000 and 2,600 m (mean \pm s.e.m., $n = 10$ per community and site), growing in sites at either 1,400 m, 2,000 m or 2,600 m.

Extended Data Table 1 | Environmental characteristics of the three study sites

Variable	1,400 m site	2,000 m site	2,600 m site
Latitude (°N)	46.8692	46.8879	46.8931
Longitude (°E)	9.4900	9.4895	9.4705
Aspect	SE	SE	SE
Daily mean temp. 2013 (°C)	6.8	4.2	1.2
Daily mean temp. 2014* (°C)	7.7	4.6	0.4
Two year average of daily mean temperature (°C)	7.0	4.4	1.0
Interpolated annual precipitation (mm)	1169	1355	1573

Temperatures were determined from temperature loggers placed at each site (see Extended Data Fig. 1). Precipitation data were obtained from interpolations of Swiss climate (ca. 1961–1990) at 50 m resolution²⁸.

*Until 2 July 2014.

Extended Data Table 2 | Characteristics of the focal species

Species	Family	Growth form	Seed mass (mg)	Inflorescence height (cm)	Terminal velocity of seeds (m/s)	Predicted maximum dispersal (m)
<i>Anthyllis alpestris</i>	Fabaceae	Erect	4.63	8.0	1.5	1.9
<i>Plantago atrata</i>	Plantaginaceae	Rosette	2.98	6.5	7.5	0.2
<i>Pulsatilla vernalis</i>	Ranunculaceae	Erect	1.73	6.0	1.0	6.9
<i>Scabiosa lucida</i>	Caprifoliaceae	Rosette	1.14	13.0	2.1	0.6

Terminal velocity was taken from the Dispersal and Diaspore database²⁹. Maximum dispersal was predicted from species' family, terminal velocity and dispersal mode (only *P. vernalis* seeds are specialized for wind dispersal), following statistical model 1 of ref. 30 implemented in R with the function 'dispeRsal' where growth form = herb for all species.

Extended Data Table 3 | Statistical analysis of the effects of soil biota on plant biomass

Species	Seed origin	<i>n</i>	χ^2	<i>P</i>
<i>Helianthemum nummularium</i>	1,400 m	20	0.88	0.349
<i>Plantago lanceolata</i>	1,400 m	16	8.13	0.004
<i>Plantago media</i>	1,400 m	20	8.27	0.004
<i>Plantago atrata</i>	2,000 m	20	6.72	0.010
<i>Pulsatilla vernalis</i>	2,000 m	12	0.37	0.544
<i>Scabiosa lucida</i>	2,000 m	18	5.38	0.020

All models contained block as a random effect; d.f. = 1 for χ^2 tests of the effects of soil biota (from 1,400 or 2,000 m above sea level) on above-ground biomass.

Extended Data Table 4 | Analysis of herbivory on four alpine plant species

Year	Species	Site	χ^2	<i>P</i>	<i>n</i>	<i>n</i> block	<i>r</i> with survival	<i>r</i> with biomass
2013	<i>Anthyllis alpestris</i>	1,400 m	4.93	0.026	31	9	-0.16	0.08
		2,000 m	0.03	0.854	59	10	-0.01	-0.27
	<i>Plantago atrata</i>	1,400 m	1.16	0.282	56	10	-0.08	-0.29
		2,000 m	0.06	0.808	59	10	-0.21	0.17
	<i>Pulsatilla vernalis</i>	1,400 m	3.03	0.082	46	10	0.13	-0.12
		2,000 m	NA*	NA	NA	NA	NA	NA
	<i>Scabiosa lucida</i>	1,400 m	4.55	0.033	49	10	-0.25	-0.08
		2,000 m	0.25	0.614	56	10	0.02	0.00
2014	<i>Anthyllis alpestris</i>	1,400 m	0.77	0.380	14	7	-0.17	0.29
		2,000 m	4.66	0.031	49	10	-0.07	0.24
	<i>Plantago atrata</i>	1,400 m	1.77	0.183	40	10	-0.08	0.02
		2,000 m	0.53	0.467	52	10	0.04	0.00
	<i>Pulsatilla vernalis</i>	1,400 m	0.05	0.824	16	8	-0.20	0.27
		2,000 m	1.86	0.173	42	10	-0.12	-0.38
	<i>Scabiosa lucida</i>	1,400 m	NA*	NA	NA	NA	NA	NA
		2,000 m	0.52	0.470	48	10	NA	0.20

The effect of competitor community identity on the incidence of herbivory on focal species was assessed with mixed-effects models fitted separately for each species and site, including plot nested in block as random effects and log(initial size) as a fixed effect. Shown are likelihood ratio tests (d.f. = 1), and the total number of observations (*n*) and experimental units (*n* block) for each model. Also shown are correlations of the incidence of herbivory with biomass and survival after 2 years (*P* < 0.05 indicated in bold). No tests are significant after Holm–Bonferroni correction.

*Model could not be fitted because herbivory was constant.

A sexually dimorphic hypothalamic circuit controls maternal care and oxytocin secretion

Niv Scott¹, Matthias Prigge¹, Ofer Yizhar¹ & Tali Kimchi¹

It is commonly assumed, but has rarely been demonstrated^{1,2}, that sex differences in behaviour arise from sexual dimorphism in the underlying neural circuits^{3,4}. Parental care is a complex stereotypic behaviour towards offspring that is shared by numerous species⁵. Mice display profound sex differences in offspring-directed behaviours. At their first encounter, virgin females behave maternally towards alien pups while males will usually ignore the pups or attack them^{6–9}. Here we show that tyrosine hydroxylase (TH)-expressing neurons in the anteroventral periventricular nucleus (AVPV) of the mouse hypothalamus are more numerous in mothers than in virgin females and males, and govern parental behaviours in a sex-specific manner. In females, ablating the AVPV TH⁺ neurons impairs maternal behaviour whereas optogenetic stimulation or increased TH expression in these cells enhance maternal care. In males, however, this same neuronal cluster has no effect on parental care but rather suppresses inter-male aggression. Furthermore, optogenetic activation or increased TH expression in the AVPV TH⁺ neurons of female mice increases circulating oxytocin, whereas their ablation reduces oxytocin levels. Finally, we show that AVPV TH⁺ neurons relay a monosynaptic input to oxytocin-expressing neurons in the paraventricular nucleus. Our findings uncover a previously unknown role for this neuronal population in the control of maternal care and oxytocin secretion, and provide evidence for a causal relationship between sexual dimorphism in the adult brain and sex differences in parental behaviour.

The hypothalamus contains several sexually dimorphic nuclei, and has a critical role in coordinating sexual dimorphism in reproductive behaviours and physiological responses to environmental cues^{4,10}. Among the sexually dimorphic hypothalamic nuclei, the AVPV (Fig. 1a) is unique as it possesses several female-biased sexually dimorphic characteristics^{2,11–13}, including a markedly larger number of tyrosine hydroxylase (TH)-immunoreactive neurons in females than in males^{4,14}. The presence of TH in these cells suggests that they are dopaminergic, but this has not been tested directly.

We first characterized the TH⁺ AVPV neurons in adult mice and found that 93% ($\pm 1\%$, $n = 6$ mice) of them co-express DOPA decarboxylase (DDC), the enzyme responsible for the formation of dopamine from DOPA (3,4-dihydroxyphenylalanine) (Fig. 1b). Since dopamine signalling is known to enhance mother–pup interactions^{6,15,16}, we hypothesized that TH⁺ AVPV neurons are part of a sexually dimorphic circuit that underlies sex differences in parental care. We first tested whether parenthood is associated with changes in the number of TH⁺ AVPV cells. In sexually naive (virgin) females, the number of TH-positive neurons was significantly higher than that in virgin males (Fig. 1c, d), consistent with previous work^{10,14}. Surprisingly, we further found that in postpartum females this number was significantly higher than in virgin females (725 ± 21 and 493 ± 60 , respectively; $P < 0.001$), whereas no such differences were observed between parental and virgin males (271 ± 21 and 273 ± 19 , respectively; Fig. 1c, d).

To assess the functional relationship between sexual dimorphism in TH⁺ AVPV neurons and sex differences in parental behaviour, we employed three complementary cell-type-specific strategies to selectively manipulate TH⁺ AVPV neurons in adult males and females: selective ablation, overexpression of TH and optogenetic activation (Fig. 2a–e). To ablate TH⁺ AVPV neurons, we bilaterally injected the neurotoxin 6-hydroxydopamine¹⁷ into the AVPV of wild-type mice. This effectively ablated the majority of TH⁺ AVPV neurons, but did not affect major TH-expressing neuronal populations in other brain regions (Fig. 2a and Extended Data Fig. 1a–d). Virgin females with bilateral ablation of TH⁺ AVPV neurons exhibited a profound reduction in maternal displays, including a significantly longer latency to retrieve foreign pups to the nest and a smaller number of pups retrieved compared to control littermates (Fig. 2f, g and Extended Data Fig. 2a, b). TH-ablated virgin females also exhibited a shorter duration of crouching over the pups and a shorter overall duration of maternal behaviour (Fig. 2h and Extended Data Fig. 2c, d). In postpartum females (mothers), TH⁺ AVPV neuronal ablation induced similar deficits in

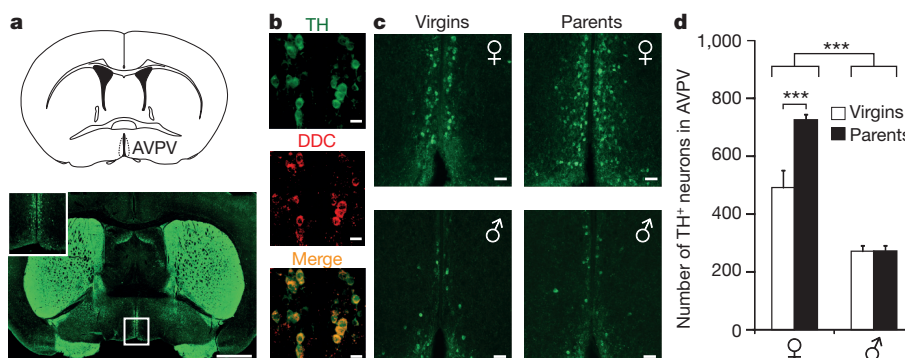


Figure 1 | TH expression in the AVPV is sexually dimorphic and enhanced in postpartum females. **a**, Schematic drawing of mouse AVPV (top) and confocal images of a coronal brain section immunostained for TH (bottom). Inset shows higher magnification image of the AVPV region. Scale bar, 1 mm. **b**, AVPV in a coronal slice from a female mouse immunostained

for TH (green) and DDC (red). Scale bars, 20 μ m. **c**, Immunostaining for TH in AVPV of female and male virgins and parents. Scale bars, 50 μ m. **d**, Total numbers of TH⁺ neurons in AVPVs of virgin females, virgin males, postpartum females and newly parental males. Data are means \pm s.e.m., $n = 5$ per group; *** $P < 0.001$; two-way ANOVA with Fisher's multiple comparisons.

¹Department of Neurobiology, Weizmann Institute of Science, 76100 Rehovot, Israel.

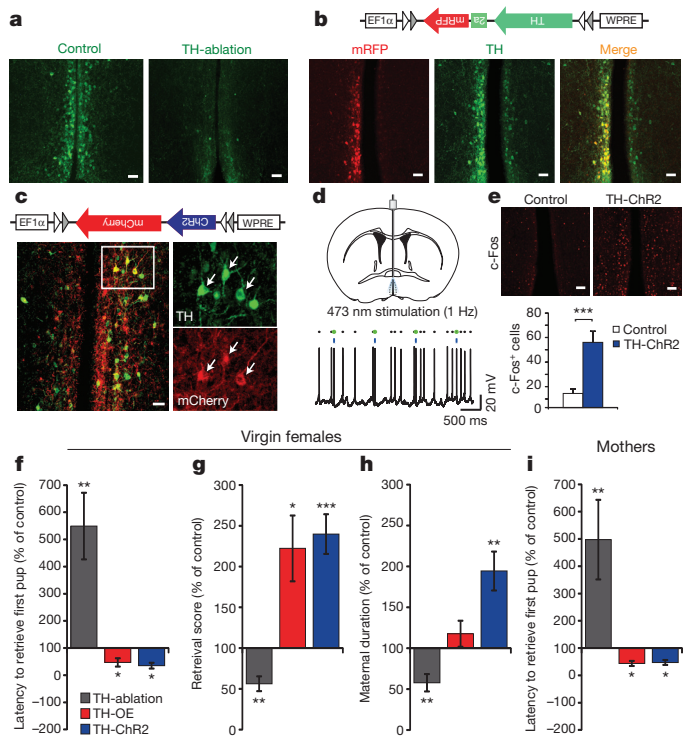


Figure 2 | TH⁺ AVPV neurons promote maternal behaviour in virgin and postpartum females. **a**, Coronal section of female AVPV immunostained for TH (green). Scale bars, 50 μ m. **b**, Schematic drawing of TH overexpression vector (top) and confocal images of coronal sections of unilateral TH-overexpression in the AVPV (bottom). mRFP labels virally transduced neurons (red). TH-immunoreactive neurons are labelled in green. Scale bars, 50 μ m. **c**, Schematic drawing of AAV vector used to express ChR2 in TH⁺ AVPV neurons (top), confocal image depicting ChR2-mCherry expression (red) and TH immunostaining (green) in AVPV neurons (bottom). Scale bar, 50 μ m. **d**, Schematic illustration of photostimulation of TH⁺ AVPV neurons expressing ChR2 (top) and an example voltage recording of a ChR2-expressing AVPV neuron stimulated with light pulses at 1 Hz (bottom). Blue bars represent 10-ms light pulses (475 nm, 19 mW mm⁻²); black dots indicate detected action potentials; green outlines represent action potentials that are time-locked to light pulses. **e**, Representative images from a TH-ChR2 mouse and TH-EYFP (control) mouse photostimulated in the AVPV. c-Fos-immunoreactive cells are labelled in red (top). Scale bars, 20 μ m. Number of c-Fos-expressing cells in AVPV following optogenetic stimulation in TH-ChR2 and control mice (bottom) ($n_{\text{ChR2}} = 6$, $n_{\text{control}} = 6$, $***P < 0.001$, two-tailed Student's *t*-test). **f–h**, Quantification of maternal behaviour of virgin females with TH-ablation (grey bars), TH-OE (red bars) and TH-ChR2 (blue bars) relative to control groups (TH-ablation, $n_{\text{ablation}} = 13$, $n_{\text{control}} = 13$; TH-OE, $n_{\text{OE}} = 9$, $n_{\text{control}} = 10$; TH-ChR2, $n_{\text{ChR2}} = 12$, $n_{\text{control}} = 14$; $*P < 0.05$, $**P < 0.01$, $***P < 0.001$, Mann–Whitney U-test). **i**, Latency to retrieve the 1st pup of postpartum females with TH-ablation, TH-OE and TH-ChR2 relative to control groups (TH-ablation, $n_{\text{ablation}} = 11$, $n_{\text{control}} = 11$; TH-OE, $n_{\text{OE}} = 8$, $n_{\text{control}} = 7$; TH-ChR2, $n_{\text{ChR2}} = 6$, $n_{\text{control}} = 6$; $*P < 0.05$, $**P < 0.01$, Mann–Whitney U-test). Data are means \pm s.e.m.

maternal behaviour, manifested as longer latencies to pup retrieval (Fig. 2i and Extended Data Fig. 2e). The AVPV was reported to play a role in the regulation of GnRH neurons (known to control ovulation)^{18,19}. However, we found that this TH⁺-specific ablation in the AVPV did not cause marked changes in female sexual behaviour, oestrous cycle and reproductive success (Extended Data Fig. 3).

The expression level of TH in some dopaminergic neurons is activity-dependent and modulated by sensory stimuli^{20,21}. To examine whether elevated TH expression levels in TH⁺ AVPV neurons are causal in driving maternal behaviour, we overexpressed TH specifically in TH⁺ AVPV neurons of TH-Cre females using a Cre-dependent TH-expressing adeno-associated virus (AAV) vector (TH-OE; Fig. 2b). In contrast

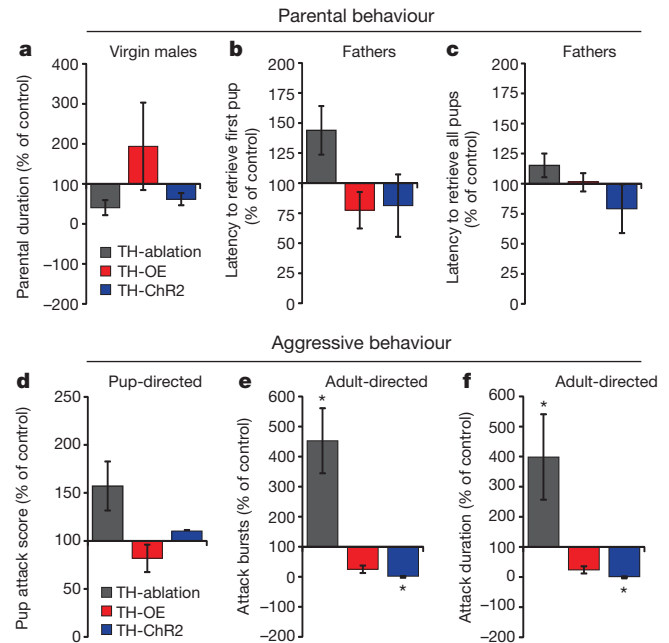


Figure 3 | TH⁺ AVPV neurons suppress inter-male aggression but do not regulate parental behaviour in males. **a, d**, Quantification of pup-directed behaviours of virgin males with TH-ablation, TH-OE and TH-ChR2 relative to control groups. **b, c**, Pup retrieval of paternal males with TH-ablation, TH-OE and TH-ChR2 relative to control groups. **e, f**, Quantification of inter-male aggression in the resident-intruder test. Data are means \pm s.e.m., TH-ablation, $n_{\text{ablation}} = 12$, $n_{\text{control}} = 12$; TH-OE, $n_{\text{OE}} = 10$, $n_{\text{control}} = 10$; TH-ChR2, $n_{\text{ChR2}} = 8$, $n_{\text{control}} = 7$; $*P < 0.05$, Mann–Whitney U-test.

to the effect of TH⁺ cell ablation, TH-OE promoted maternal behaviour in both virgin and postpartum females (Fig. 2f–i). TH-OE virgin females presented significantly shorter latencies to pup retrieval, and retrieved more pups to the nest compared to virgin females injected with a control mCherry-expressing AAV vector (Fig. 2f, g and Extended Data Fig. 4a, b). Similarly, TH-OE in mothers induced shorter latencies to pup retrieval (Fig. 2i and Extended Data Fig. 4e). This suggests that maternal pup retrieval is facilitated by increased TH levels in TH⁺ AVPV cells.

To determine whether activation of the TH⁺ AVPV circuit can promote maternal behaviour, we injected virgin TH-Cre females with a Cre-dependent channelrhodopsin-2 AAV vector into the AVPV (TH-ChR2; Fig. 2c). We first recorded in acute brain slices from ChR2⁺ and ChR2⁻ neurons in the AVPV to verify effective photostimulation of ChR2-expressing cells. The spontaneous firing rates of TH⁺ AVPV neurons were similar to those described previously for AVPV neurons²², and their intrinsic electrophysiological properties resembled those of midbrain dopaminergic neurons²³ (Fig. 2d and Extended Data Fig. 5a–j). We also confirmed *in vivo* activation of these neurons using c-Fos staining (Fig. 2e). Tonic 1 Hz photostimulation of TH⁺ AVPV neurons in virgin females elicited a shorter latency to retrieve pups to the nest and increased the number of pups retrieved (Fig. 2f, g; Extended Data Fig. 6a, b and Supplementary Videos 1, 2). Moreover, photostimulation promoted crouching behaviour and prolonged the overall duration of maternal care (Fig. 2h and Extended Data Fig. 6c, d). Similarly, activation of TH⁺ AVPV neurons in mothers promoted pup retrieval to the nest, but did not affect maternal aggression, an adult-directed aggressive behaviour in lactating dams (Fig. 2i and Extended Data Fig. 6e–g). In all of the TH-manipulated groups, we verified that the observed effects on maternal behaviour are not accompanied by changes in anxiety, exploration, or locomotor activity (Extended Data Fig. 7a–l). These findings indicate that TH⁺ AVPV neurons are required for the control of maternal behaviour in virgin and postpartum females.

We next examined whether TH⁺ AVPV neurons are involved in the regulation of parental behaviour in males. To address this, we manipulated TH⁺ AVPV neurons in males using the same TH-specific strategies. In contrast to the robust effects seen in females, neither ablation nor TH-overexpression or optogenetic activation of TH⁺ AVPV neurons induced significant changes in parental behaviour in virgin and parental (fathers) males (Fig. 3a–c, Extended Data Figs 2f–h, 4f–h and 6h–j). Notably, however, ablation of TH⁺ AVPV neurons in virgin males induced a trend towards increased pup-directed aggression ($P = 0.09$; Fig. 3d and Extended Data Fig. 2h). To further assess whether TH⁺ AVPV cells contribute to the regulation of male aggression, we tested the effects of our TH-manipulations on inter-male aggression by using the resident–intruder assay²⁴. We found that males in which TH⁺ AVPV neurons were ablated presented an increase in attack bursts and in the total duration of aggression (Fig. 3e, f and Extended Data Fig. 2i, j). In contrast, optogenetic activation of TH⁺ AVPV neurons led to a significant reduction in these behavioural measures (Fig. 3e, f; Extended Data Fig. 6k, l and Supplementary Videos 3, 4). TH-manipulated males displayed no changes in anxiety, exploration or locomotor activity (Extended Data Fig. 8a–j). We conclude that in males, TH⁺ AVPV neurons act as negative regulators of inter-male aggression but do not contribute to regulation of parental care.

Several hormones, including oestradiol, corticosterone, prolactin and oxytocin, have been implicated in the regulation of sex-specific parental behaviour^{16,25}. This could suggest that the regulation of maternal behaviour by TH⁺ AVPV neurons is associated with the action of these hormones. We measured changes in the peripheral levels of these hormones in the absence of pups (Fig. 4a and Extended Data Fig. 9a–e). We found that oxytocin (OT) levels, but not the levels of other hormones, were reduced following TH⁺ AVPV neuronal ablation and elevated in TH-OE virgin females compared to controls (Fig. 4a and Extended Data Fig. 9b, d, e). Remarkably, we further found that optogenetic stimulation of TH⁺ AVPV neurons at 1 Hz for 10 min was sufficient to induce a significant increase in OT levels in Chr2-expressing virgin females (Fig. 4a). In contrast, none of the TH⁺ AVPV manipulations induced significant changes in circulating OT levels in virgin males (Extended Data Fig. 9a). TH⁺ AVPV neurons thus seem to directly regulate the release of circulating OT in females.

We next explored the anatomical and functional connectivity between TH⁺ AVPV neurons and OT-secreting neurons in the paraventricular nucleus (PVN) and the supraoptic nucleus (SON)²⁶. We fluorescently labelled the axonal projections of the TH⁺ AVPV neurons by injecting a Cre-dependent enhanced yellow fluorescent protein (EYFP)-expressing viral vector into the AVPV of TH-Cre mice. Projection analysis revealed several brain regions containing fluorescently labelled TH⁺ fibres in both males and females (Extended Data Fig. 10a–d). Among the brain regions that showed the densest projections were the medial preoptic area, a region associated with parental behaviour^{6,27,28}, and the PVN. Immunostaining for OT in the PVN revealed dense TH⁺ AVPV fibres in close proximity to OT⁺ PVN neurons (Fig. 4b). Additionally, we injected a Cre-dependent anterograde viral tracer (H129ΔTK-TT) into the AVPV of TH-Cre females. Immunohistochemical staining for OT in the PVN showed that $55 \pm 7\%$ of OT⁺ neurons co-labelled with the anterograde virus (Fig. 4c, e). In contrast, similar staining in the SON did not reveal any virally-transduced OT⁺ neurons (Fig. 4d, e).

To test for a functional synaptic connection between TH⁺ AVPV neurons and OT⁺ PVN cells, we obtained whole-cell patch clamp recordings from PVN OT⁺ cells and photostimulated Chr2⁺ axonal projections from the AVPV. For these recordings, we labelled OT⁺ PVN neurons of TH-Cre females with an AAV vector encoding the Venus fluorophore under an OT promoter. A second AAV, encoding a Cre-dependent Chr2–mCherry fusion protein, was injected into the AVPV. Acute slices prepared from these mice showed Chr2–mCherry-labelled fibres throughout the region containing

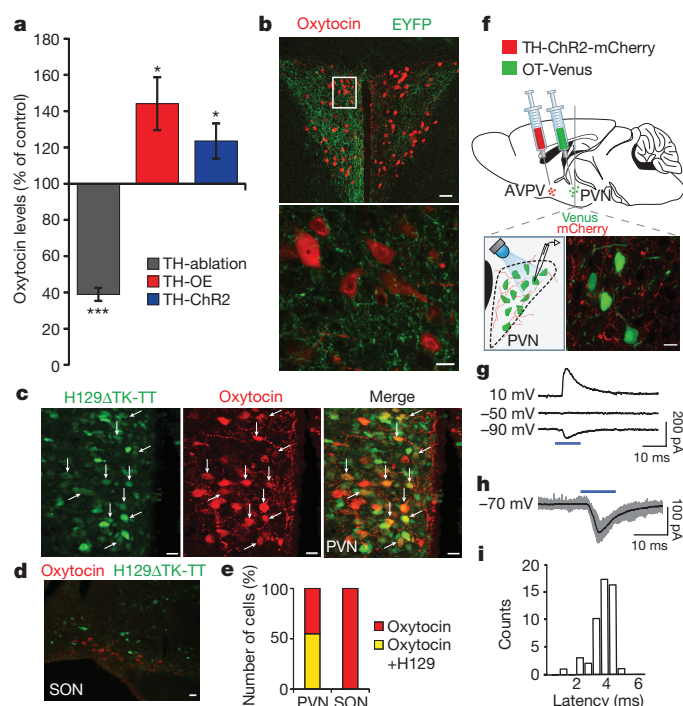


Figure 4 | TH⁺ AVPV neurons form synaptic connections with OT⁺ PVN neurons and regulate OT secretion in females. **a**, Plasma OT levels in TH-ablation, TH-OE and TH-ChR2 females relative to control groups (data are means \pm s.e.m., TH-ablation, $n_{\text{ablation}} = 8$, $n_{\text{control}} = 5$; TH-OE, $n_{\text{OE}} = 9$, $n_{\text{control}} = 9$; TH-ChR2, $n_{\text{Chr2}} = 7$, $n_{\text{control}} = 8$; $*P < 0.05$, $***P < 0.001$, two-tailed Student's *t*-test). **b**, Coronal sections of the PVN in TH-Cre female injected unilaterally with Cre-dependent EYFP vector into the AVPV. Anterograde fibres projecting from TH⁺ AVPV neurons (green) lie in the vicinity of OT⁺ neurons in the PVN (red). Scale bars, 50 μ m (top) and 10 μ m (bottom). **c**, Coronal sections of TH-Cre female injected with Cre-dependent anterograde tracer (H129ΔTK-TT) into the AVPV, showing virally labelled OT⁺ neurons in the PVN. Arrows indicate virally transduced neurons (green pseudocolour) that are OT⁺ (red pseudocolour). Scale bars, 10 μ m. **d**, Coronal section of the SON of TH-Cre female injected with Cre-dependent anterograde tracer into the AVPV, showing virally labelled neurons (green) and OT⁺ neurons (red). Scale bar, 10 μ m. **e**, Percentage of virally transduced OT⁺ neurons in the PVN and SON (PVN: $55 \pm 7\%$, SON: none, $n = 3$ mice). **f**, Illustration of the experimental approach. Top, schematic drawing of a mouse brain in a sagittal view injected with AAV-DIO-ChR2–mCherry into the AVPV and with AAV-OT–Venus into the PVN. Bottom, schematic drawing and confocal image of PVN coronal section showing the OT⁺ neurons transduced with AAV-OT–Venus (green) and fibres from TH⁺ AVPV neurons expressing Chr2–mCherry (red). Scale bar, 10 μ m. **g**, Light-evoked responses in OT⁺ cells recorded in voltage clamp mode at the specified holding potentials. **h**, Average light response (black) overlays 50 single-trial responses (grey) of a representative OT⁺ cell clamped to -70 mV. **i**, Distribution of light-induced response latencies from 50 light stimuli recorded in a single OT⁺ cell (median = 3.77 ms from light onset). SON, supraoptic nucleus; PVN, paraventricular nucleus.

OT⁺ neurons, confirming the results of our viral tracing experiments (Fig. 4f). Photostimulation of TH-ChR2 fibres in the PVN evoked fast excitatory post-synaptic currents in 7 out of 19 OT⁺ PVN neurons recorded (3.81 ± 0.83 ms latency from blue light onset; $n = 7$ cells from 3 mice; Fig. 4g–i). These results indicate that a monosynaptic connection exists between the TH⁺ AVPV neurons and OT⁺ PVN cells, suggesting that TH⁺ AVPV neurons can facilitate OT release from OT⁺ PVN neurons.

Our study reveals that a female-biased cluster of TH-expressing neurons in the AVPV has an essential role in the control of sex-specific behaviours in both sexes. Whereas in females this neural population acts to promote maternal care and OT secretion, in males it is rather involved in the suppression of adult-directed aggression. Moreover, we

have uncovered a novel monosynaptic circuit linking the TH⁺ AVPV neurons with PVN OT⁺ neurons. In rodents, as in many mammals, OT signalling is necessary for lactation and has been implicated as a central neuroendocrine mediator of maternal behaviour^{5,25,29,30}. Based on our findings and on these previous reports, we propose that in females, pup-mediated signals trigger the activation of TH⁺ AVPV neurons, which in turn activate oxytocinergic neurons in the PVN, releasing OT into the brain and periphery and facilitating instinctive parental behaviours (Extended Data Fig. 10e).

Finally, although manipulation of TH⁺ AVPV neurons markedly altered sex-specific behaviours in both males and females, the behaviours displayed by manipulated animals of both sexes did not exceed the boundaries of sex-typical behaviours. These data highlight the critical role of intrinsic sex differences in the brain in setting the distinct behavioural repertoire displayed by males and females. Within these boundaries, changes in the activity of sexually-dimorphic neural circuits allow dynamic modulation of adaptive behavioural responses to sex-specific challenges.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 22 March; accepted 31 July 2015.

Published online 16 September 2015.

- Xu, X. *et al.* Modular genetic control of sexually dimorphic behaviors. *Cell* **148**, 596–607 (2012).
- Yang, C. F. *et al.* Sexually dimorphic neurons in the ventromedial hypothalamus govern mating in both sexes and aggression in males. *Cell* **153**, 896–909 (2013).
- De Vries, G. J. & Villalba, C. Brain sexual dimorphism and sex differences in parental and other social behaviors. *Ann. NY Acad. Sci.* **807**, 273–286 (1997).
- Simerly, R. B. Wired for reproduction: organization and development of sexually dimorphic circuits in the mammalian forebrain. *Annu. Rev. Neurosci.* **25**, 507–536 (2002).
- Dulac, C., O'Connell, L. A. & Wu, Z. Neural control of maternal and paternal behaviors. *Science* **345**, 765–770 (2014).
- Kuroda, K. O., Tachikawa, K., Yoshida, S., Tsuneoka, Y. & Numan, M. Neuromolecular basis of parental behavior in laboratory mice and rats: with special emphasis on technical issues of using mouse genetics. *Prog. Neuropsychopharmacol. Biol. Psychiatry* **35**, 1205–1231 (2011).
- Vom Saal, F. S. Time-contingent change in infanticide and parental behavior induced by ejaculation in male mice. *Physiol. Behav.* **34**, 7–15 (1985).
- Chalfin, L. *et al.* Mapping ecologically relevant social behaviours by gene knockout in wild mice. *Nat. Commun.* **5**, 4569 (2014).
- Lonstein, J. S. & De Vries, G. J. Sex differences in the parental behavior of rodents. *Neurosci. Biobehav. Rev.* **24**, 669–686 (2000).
- Semaan, S. J. & Kauffman, A. S. Sexual differentiation and development of forebrain reproductive circuits. *Curr. Opin. Neurobiol.* **20**, 424–431 (2010).
- Simerly, R. B. Organization and regulation of sexually dimorphic neuroendocrine pathways. *Behav. Brain Res.* **92**, 195–203 (1998).
- Cao, J. & Patisaul, H. B. Sexually dimorphic expression of hypothalamic estrogen receptors α and β and Kiss1 in neonatal male and female rats. *J. Comp. Neurol.* **519**, 2954–2977 (2011).
- Zuloaga, D. G., Zuloaga, K. L., Hinds, L. R., Carbone, D. L. & Handa, R. J. Estrogen receptor β expression in the mouse forebrain: age and sex differences. *J. Comp. Neurol.* **522**, 358–371 (2014).
- Simerly, R. B., Zee, M. C., Pendleton, J. W., Lubahn, D. B. & Korach, K. S. Estrogen receptor-dependent sexual differentiation of dopaminergic neurons in the preoptic region of the mouse. *Proc. Natl Acad. Sci. USA* **94**, 14077–14082 (1997).
- Love, T. M. Oxytocin, motivation and the role of dopamine. *Pharmacol. Biochem. Behav.* **119**, 49–60 (2014).
- Lonstein, J. S. & Morrell, J. I. in *Handbook of Neurochemistry and Molecular Neurobiology: Behavioral Neurochemistry, Neuroendocrinology and Molecular Neurobiology* (eds Lejtha, A. & Blaustein J.) Ch. 5, 197–229 (Springer-Verlag, 2007).
- Iancu, R., Mohapel, P., Brundin, P. & Paul, G. Behavioral characterization of a unilateral 6-OHDA-lesion model of Parkinson's disease in mice. *Behav. Brain Res.* **162**, 1–10 (2005).
- Liu, X. & Herbison, A. E. Dopamine regulation of gonadotropin-releasing hormone neuron excitability in male and female mice. *Endocrinology* **154**, 340–350 (2013).
- Han, S. K. *et al.* Activation of gonadotropin-releasing hormone neurons by kisspeptin as a neuroendocrine switch for the onset of puberty. *J. Neurosci.* **25**, 11349–11356 (2005).
- Spitzer, N. C. Activity-dependent neurotransmitter respecification. *Nature Rev. Neurosci.* **13**, 94–106 (2012).
- Aumann, T. D., Tomas, D. & Horne, M. K. Environmental and behavioral modulation of the number of substantia nigra dopamine neurons in adult mice. *Brain Behav.* **3**, 617–625 (2013).
- Ducet, E., Gaidamaka, G. & Herbison, A. E. Electrical and morphological characteristics of anteroventral periventricular nucleus kisspeptin and other neurons in the female mouse. *Endocrinology* **151**, 2223–2232 (2010).
- Margolis, E. B., Lock, H., Hjelmstad, G. O. & Fields, H. L. The ventral tegmental area revisited: is there an electrophysiological marker for dopaminergic neurons? *J. Physiol. (Lond.)* **577**, 907–924 (2006).
- Kimchi, T., Xu, J. & Dulac, C. A functional circuit underlying male sexual behaviour in the female mouse brain. *Nature* **448**, 1009–1014 (2007).
- Rilling, J. K. & Young, L. J. The biology of mammalian parenting and its effect on offspring social development. *Science* **345**, 771–776 (2014).
- Lee, H.-J., Macbeth, A. H., Pagani, J. H. & Young, W. S. III. Oxytocin: the great facilitator of life. *Prog. Neurobiol.* **88**, 127–151 (2009).
- Wu, Z., Autry, A. E., Bergan, J. F., Watabe-Uchida, M. & Dulac, C. G. Galanin neurons in the medial preoptic area govern parental behaviour. *Nature* **509**, 325–330 (2014).
- Numan, M. & Woodside, B. Maternity: neural mechanisms, motivational processes, and physiological adaptations. *Behav. Neurosci.* **124**, 715–741 (2010).
- Numan, M. & Insel, T. R. *The Neurobiology of Parental Behavior* (Springer-Verlag, 2003).
- Marlin, B. J., Mitre, M., D'amour, J. A., Chao, M. V. & Froemke, R. C. Oxytocin enables maternal behaviour by balancing cortical inhibition. *Nature* **520**, 499–504 (2015).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank M. Dayan, E. Elharrar, I. Sofer, Y. Pen, Y. Beny, N. Zhilka and E. Massasa for their assistance with the experiments; R. Levy for help with virus production; A. Chen and A. Ramot for sharing equipment and reagents. A. Mizrahi, S. Wagner, and the Yizhar and Kimchi groups, for comments on the manuscript; D. Anderson for providing the Cre-dependent anterograde virus and V. Grinevich for providing the OT:Venus virus. M.P. was supported by the Clore Center for Biological Physics and a Minerva postdoctoral fellowship. This work was supported by grants from Minerva Foundation 711131, Women's Health Research Center, Gruber Foundation 720667, ISF 1324/15, the Jenna and Julia Birnbach Career Development Chair to T.K.; and Marie Curie CIG 321919, ERC StG 337637, ISF 1351/12, the Nollman Career Development Chair to O.Y.

Author Contributions N.S. performed all behavioural, neuronal tracing and *in vivo* physiological experiments. M.P. cloned the TH-overexpression viral vector and performed the *in vitro* electrophysiology experiments. O.Y. supervised the electrophysiology experiments and assisted in designing the optogenetic assays. T.K. planned and supervised all the experiments. T.K. and N.S. interpreted the results and wrote the paper together with M.P. and O.Y.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to T.K. (tali.kimchi@weizmann.ac.il) or O.Y. (ofer.yizhar@weizmann.ac.il)

METHODS

Animals. Male and female mice (12 to 20 weeks old) were used in all experiments. For TH-ablation experiments wild-type mice (CD-1-Harlan Laboratory) were used. For TH-overexpression, TH⁺ optogenetic stimulation and tracing experiments, TH-IRES-Cre³¹ male and female mice were used. Littermates were randomly assigned to experiment or control groups. We confirmed the specificity of Cre-expressing neurons in virgin females by showing that a majority of Cre-expressing cells overlap with TH-immunoreactive cells in the AVPV (EYFP (Cre) cells expressing TH = 80 ± 3%; TH cells expressing EYFP (Cre) = 91 ± 3%, *n* = 6 mice). Compared with reports of reduced specificity of this mouse line to dopamine neurons in the ventral tegmental area³², these data indicate that specificity in the AVPV is relatively high. Furthermore, increased TH expression in maternal females might further increase the specificity values. All mice were bred and housed in a specific-pathogen-free animal facility; they were maintained on a reverse 12:12 h light–dark cycle with food and water *ad libitum*. All experimental procedures were approved by the Institutional Animal Care and Use Committee (IACUC) at the Weizmann Institute of Science.

Viral vectors. pAAV5-Efl α -DIO-ChR2(E123T/T159C)-mCherry, pAAV5-Efl α -DIO-EYFP and pAAV5-Efl α -DIO-mCherry were acquired from UNC Gene Therapy Center. H129- Δ TK-TT virus³³ was a gift of D. J. Anderson and AAV-OTpr-Venus virus³⁴ was a gift of V. Grinevich. A Cre-dependent vector for TH-overexpression (pAAV-Efl α -DIO-TH-p2A-mRFP) was produced using overlap extension PCR. A 2A peptide was inserted between TH and mRFP coding sequences to allow unperturbed function of TH. The forward and reverse primer for the entire PCR fragment contained an *Asc*I and *Nhe*I restriction sites. The resulting amplicon was digested and cloned into the pAAV-Efl α -DIO backbone³⁵. **6-OHDA.** 6-hydroxydopamine hydrochloride (Sigma-Aldrich) was dissolved in 0.1% ascorbic acid saline solution to final concentration of 6 μ g μ l⁻¹ and prepared fresh every day before injections into the AVPV.

Surgery. Mice were anaesthetized with isoflurane and mounted on a stereotaxic frame (myNeuroLab). Virus or agent was bilaterally or unilaterally injected into the AVPV (for behaviour or tracing, respectively) using a Hamilton syringe at an injection rate of 0.05–0.15 μ l min⁻¹. AVPV injection coordinates were anteroposterior (AP): 0.25 mm, mediolateral (ML): \pm 0.15 mm, dorsoventral (DV): –5.45 mm. Injection volumes were: 0.2 μ l for TH-OE and ChR2 viruses; 0.06 μ l and 0.1 μ l for AAV-EYFP and H129 Δ TK-TT tracing experiments, respectively; 1 μ l for 6-OHDA. AAV-OTpr-Venus was bilaterally injected into the PVN (AP: –0.8 mm, ML: \pm 0.2, DV: –4.75) at a volume of 0.3 μ l. ChR2 animals used for behavioural testing were implanted with 200 μ m fibre-optic cannulae (Thorlabs). Fibres were located medially and above the AVPV of both hemispheres (Fig. 2d). The cannula was secured using dental cement. Mice were left to recover for at least 3 weeks before behavioural assays.

Behavioural assays. All behavioural tests were performed during the dark phase and under dim red light. Behaviours were recorded using a digital video recording unit and scored using Observer XT or EthoVision softwares (Noldus Information Technology) by an individual blind to the identity of the manipulated groups.

Parental behaviour. Males and females were tested for pup-directed behaviours using the exact same behavioural assay procedure. Mice were individually housed with nesting material (cotton) 24 h before the trial. Three newborn alien pups (1–3 days old, CD-1, Harlan Laboratories) were placed on the opposite side to the sleeping nest in the resident mouse's home cage and behaviour was recorded for 15 min. Females were introduced to pups for 3 consecutive days and parental behaviours were scored on day 3 when animals exhibited the highest level of parental behaviour. Pup-directed behaviours of males were scored on day 1 since their behaviour did not change with repeated exposure to pups. The scored behavioural parameters were (Supplementary Videos 1, 2): pup retrieval, carrying the pups to the cotton nest; crouching, covering the pups to maintain body temperature and (in females) lactating behaviour; nesting, building a nest around the pups; licking, licking the pups; attack, pup-directed attack; ignore, failing to approach or make contact with the pups. Parental duration was taken as the sum of all parental behaviour durations in the assay: crouching, nesting and licking. Pup retrieval was scored in points and quantified as a cumulative retrieval score in which each successful pup retrieval to the nest contributed 1 point. Pup-directed aggression (attack score) was quantified similarly, with each pup attack contributing 1 point. Parental behaviours in postpartum females were tested on postpartum day 4. For this assay, first all newborn pups were removed from the home cage of each female mouse, and then after 10 min, three of the female's own pups were placed back in the resident's home cage on the opposite side to the breeding nest for 10 min. Latencies to retrieval of the pups back to the nest were scored as described above. Fathers' parental behaviour was measured in males that were housed together with females for 21–24 days (that is, separated from the lactating females and their own pups 1–3 days after the delivery of the pups).

Maternal aggression. Postpartum females were tested for maternal aggression 8 days post-delivery. All pups were removed from cage and adult intruder male was immediately introduced to female cage for 10 min assay. Females were scored for the duration and number of their aggressive attacks towards the intruder male.

Female sexual behaviour. For 4 consecutive days oestrus cycle state was determined (see below). Only females confirmed to be in oestrus were introduced with an experienced male for 15 min assay. Female receptivity was scored as lordosis, successful male mountings and male rejections.

Male–male interactions. Males were separated to their home cages and introduced to 6 to 7 week-old C57Bl/6 unfamiliar male intruders for 10 min. The male intruders were swabbed with 60 μ l urine collected from sexually mature and experienced male mice. All experiments were videotaped and the behaviours of the resident animals were scored for the following parameters: number of attacks, latency to attack and total duration of aggression (Supplementary Videos 3, 4).

Open-field assay. The test was performed in a white Plexiglas box (50 \times 50 \times 40 cm) with an overhead lamp directed to the centre of the field, providing 120 lx of illumination on the floor. Each mouse was placed in the corner of the apparatus and its locomotion pattern was recorded for 10 min.

Elevated plus maze assay. The test was performed using a polyvinyl chloride maze comprising a central part (5 \times 5 cm), two opposing open arms (30.5 \times 5 cm), and two opposing closed arms (30.5 \times 5 \times 15 cm). The apparatus was raised to a height of 53.5 cm, and the open arms were provided with 6 lx of illumination. Each mouse was placed in the centre facing an open arm and its locomotion was recorded for 5 min. The distance travelled and time spent in the open arms was measured.

ChR2 mediated *in vivo* photostimulation. Prior to all behavioural tests, animals were connected to optical fibres with a 200- μ m silica core (BFL37-200; Thorlabs) for 5–10 min habituation. To enable free movement during the test, we connected the optical fibres to an optical rotary joint (Doric Lenses QC, Canada). Light stimulation at 473 nm was provided by a DPSS laser (CrystaLaser) and lasted for the entire duration of the assay. Photostimulation of TH⁺ AVPV neurons was most effective at low frequencies (<10 Hz) and its efficacy declined at higher frequencies (Extended Data Fig. 5b). Photostimulation parameters in all assays were 1 Hz frequency, 10 ms pulse-width and 95 mW mm⁻² light power density. Laser pulses were driven by a 33220A Function Waveform Generator (Agilent Technologies, Israel). For c-Fos verification, mice were stimulated with a train of 473 nm light (5 Hz, 10 ms) for 30 min. Mice were euthanized 60 min post-stimulation and brains were processed for immunohistochemistry. The activation frequency of 5 Hz was chosen as we observed efficient excitation of the neurons using this frequency in the acute slice preparation.

Slice electrophysiology recording. Four to eight weeks after virus injection, 300- μ m-thick coronal sections of the AVPV and PVN were prepared using a vibratome (Leica TV1200S) in ice-cold sucrose cutting solution (in mM: 11 D-glucose, 234 sucrose, 2.5 KCl, 1.25 NaH₂PO₄, 10 MgSO₄, 0.5 CaCl₂, 26 NaHCO₃) oxygenated with 95% O₂/5% CO₂. Slices were then incubated at 32 °C for 30 min in high-osmolarity artificial cerebrospinal fluid (aCSF; in mM: 3.24 KCl, 11.88 glucose, 132.8 NaCl, 28.1 NaHCO₃, 1.35 NaH₂PO₄, 1.08 MgCl₂, 2.16 CaCl₂; 320 mOsm kg⁻¹, aerated with 95% O₂/5% CO₂) and another 30 min in iso-osmotic aCSF (in mM: 3 KCl, 11 glucose, 123 NaCl, 26 NaHCO₃, 1.25 NaH₂PO₄, 1 MgCl₂, 2 CaCl₂; 300 mOsm kg⁻¹, aerated with 95% O₂/5% CO₂) at 32 °C. Following recovery, slices were kept at room temperature until use. The recording chamber was perfused with oxygenated aCSF at a rate of 1.5–2 ml min⁻¹ and maintained at 32 °C. Borosilicate glass pipettes (Sutter Instrument BF100-58-10) with resistances ranging from 4–6 M Ω were pulled using a laser micropipette puller (Sutter Instrument Model P-2000) and filled with intracellular solution (in mM: 135 K-gluconate, 4 KCl, 2 NaCl, 10 HEPES, 4 EGTA, 4 MgATP, 0.3 NaTRIS, 280 mOsm kg⁻¹, pH adjusted to 7.3 with KOH). Recordings in PVN OT cells were performed with an intracellular solution (in mM: 120 Cs-gluconate, 11 CsCl, 1 MgCl₂, 1 CaCl₂, 10 HEPES, 11 EGTA, 5 QX-314, 280 mOsm kg⁻¹, pH adjusted to 7.3 with CsOH). Neurons were patched under visual guidance using infrared differential interference contrast (DIC) microscopy (Olympus BX51WIF) and an Andor Clara CCD camera. OT cells were identified based on Venus fluorescence. Recordings were carried out using a Multiclamp 700B amplifier (Axon Instruments). Optical activation of ChR2-expressing neurons was performed using 475/28 nm and 10 ms light pulses at 19 mW mm⁻² (Lumencor Spectra-X) delivered through the microscope illumination path.

Oestrus cycle analysis. Oestrous cycle stage of females was determined for 9 consecutive days. Each day at 10am a vaginal smear was collected from all females, stained with Dip Quick stain kit (Jorgensen Laboratories, Inc.) and analysed for oestrus under a light microscope. The stage of the oestrus cycle was determined as previously described³⁶.

Hormone level analysis. Upon completion of behavioural assays, mice were anaesthetized with isoflurane and blood samples were collected from the orbital

sinus. The blood was centrifuged and the supernatant was collected and stored at -80°C . Plasma oxytocin, corticosterone, oestradiol and prolactin levels were measured using ELISA kits according to the manufacturers' protocols (Enzo Life Sciences - ADI-900-153 for oxytocin, Cayman Chemical Company - 500655 for corticosterone and 582251 for oestradiol, abcam - ab100736 for prolactin).

Immunohistochemistry. Following behavioural assays and blood collection the mice were euthanized, perfused with 4% paraformaldehyde (PFA), and their brains were sectioned with a vibratome (Leica Microsystems) into $30\text{--}50\text{ }\mu\text{m}$ coronal slices. Floating brain slices were collected, washed three times in PBS, and immunostained for tyrosine hydroxylase (TH), DOPA decarboxylase (DDC), GFP, oxytocin (OT) or c-Fos using the following protocol. For $24\text{--}48\text{ h}$ at 4°C , slices were incubated in blocking solution: 10% to 20% normal human serum (NHS), 0.04% Triton; carrier: 1% NHS, 0.03% Triton; primary antibodies: rabbit/sheep anti-TH (1:1,000, Millipore), rabbit anti-DDC (1:1,000, Novus), goat anti-GFP biotinylated (1:200–1,000, Abcam), rabbit anti-c-Fos (1:1,000, Santa Cruz) and guinea-pig/rabbit anti-OT (1:1,000, Peninsula Laboratories LLC) for $24\text{--}48\text{ h}$ at 4°C ; secondary antibodies: Cy3-goat anti-rabbit, Cy5-donkey anti-guinea pig (1:200, Jackson ImmunoResearch), Alexa Flour 488-conjugated streptavidin, Alexa Flour 488/594 anti-rabbit, Alexa Flour 488/594 anti-sheep (1:200, Molecular Probes).

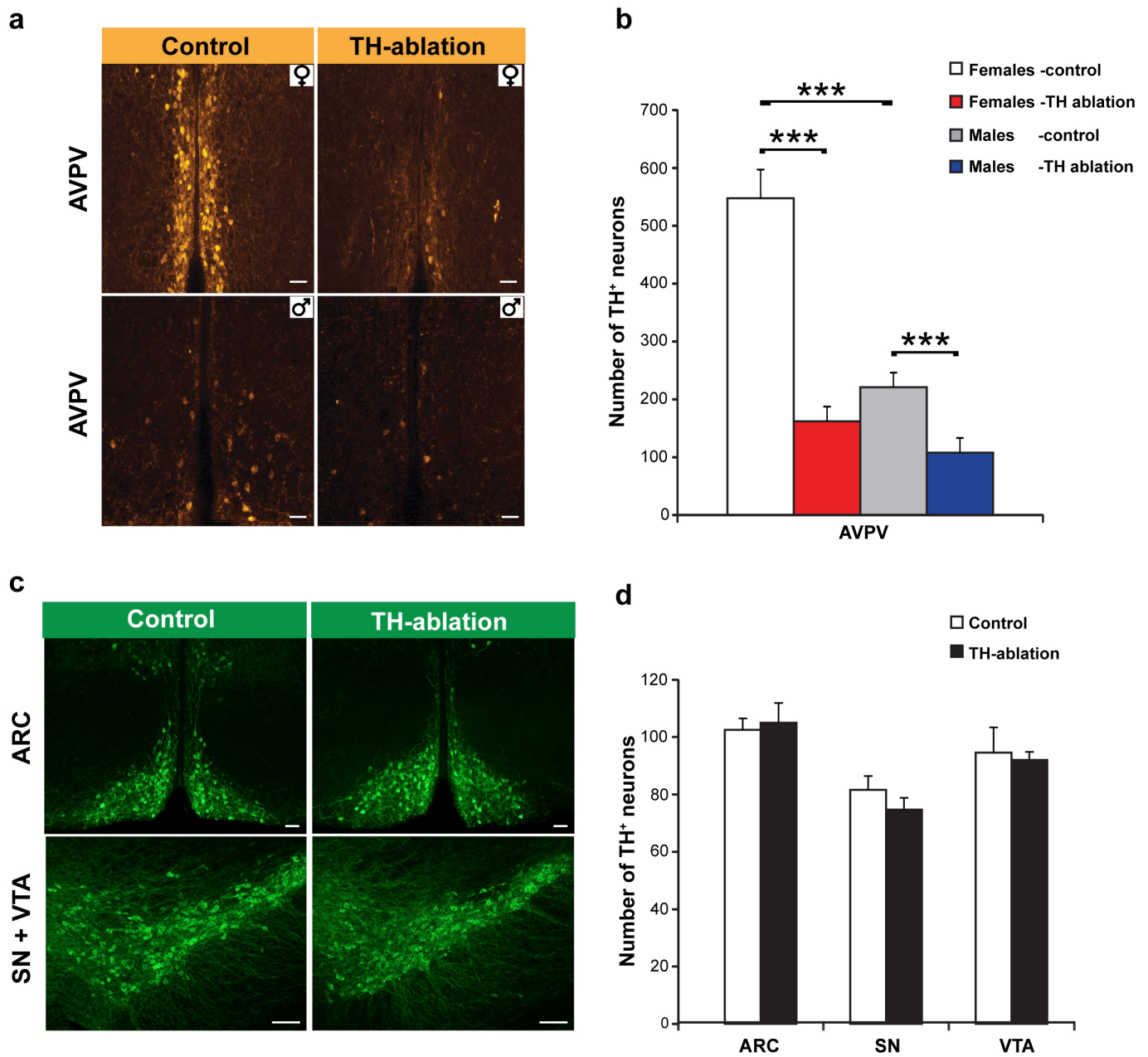
Image analysis and cell counting. All brain slices were imaged by epifluorescence microscopy (Nikon, Eclipse 80i) or by confocal microscopy (Zeiss, LSM 710) for subsequent analysis. Brain areas were determined according to their anatomy using Franklin and Paxinos Brain Atlas. For AVPV TH⁺ cell counts the entire AVPV was sliced, stained and counted. For c-Fos, anterograde tracing, viral infection and 6-OHDA lesion specificity experiments, cell counts were performed on selected brain slices, chosen in each animal according to standard anatomical markers. All counts were done manually by experimenter blind to test conditions.

Anterograde tracing. Mice injected unilaterally with AAV-DIO-EYFP virus were perfused 5 weeks post-injection, and their brains were sectioned coronally ($50\text{ }\mu\text{m}$ slices) and mounted on slides. Each brain area in which EYFP-immunoreactive

fibres were detected was scanned by confocal microscope under uniform imaging settings (LSM 710, Zeiss, Germany). Representative brain images were analysed for projection intensity using ImageJ software. For trans-synaptic anterograde tracing, females injected unilaterally with H129 Δ TK-TT were perfused 5 to 7 days post-injection as previously described²⁸. The brains were sectioned into $50\text{ }\mu\text{m}$ coronal slices, and these were mounted on slides and analysed to verify the injection site. Brain slices from the PVN and the SON were immunostained with anti-oxytocin antibody and imaged by confocal microscopy for further analysis of colocalization with dtTomato-labelled virally transduced neurons.

Statistical analysis. Sample size was determined according to the accepted practice for behavioural assays, but no statistical methods were used to predetermine sample size. All data are expressed as means \pm s.e.m. Behavioural assays were analysed using the non-parametric Mann–Whitney U-tests. Two-way ANOVA with post-hoc tests or two-tailed student's *t*-tests were used for statistical evaluation of immunostaining assays, tracing experiments and hormone analyses. Outlier samples (more than 2S.D. from the mean in >2 parameters) were excluded from experiment. All statistical tests were performed using STATISTICA software (StatSoft, Tulsa, OK).

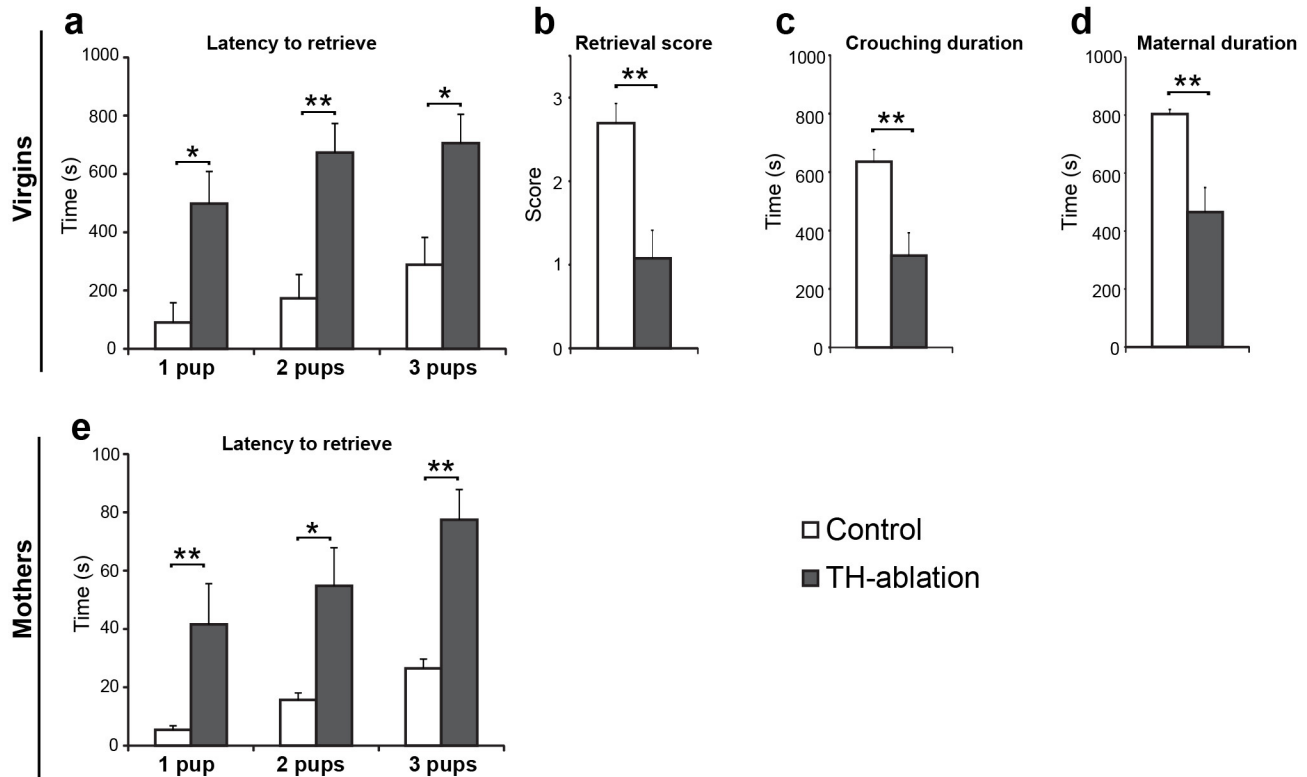
31. Lindeberg, J. *et al.* Transgenic expression of Cre recombinase from the tyrosine hydroxylase locus. *Genesis* **40**, 67–73 (2004).
32. Lammel, S. *et al.* Diversity of transgenic mouse models for selective targeting of midbrain dopamine neurons. *Neuron* **85**, 429–438 (2015).
33. Lo, L. & Anderson, D. J. A Cre-dependent, anterograde transsynaptic viral tracer for mapping output pathways of genetically marked neurons. *Neuron* **72**, 938–950 (2011).
34. Knobloch, H. S. *et al.* Evoked axonal oxytocin release in the central amygdala attenuates fear response. *Neuron* **73**, 553–566 (2012).
35. Sohal, V. S., Zhang, F., Yizhar, O. & Deisseroth, K. Parvalbumin neurons and gamma rhythms enhance cortical circuit performance. *Nature* **459**, 698–702 (2009).
36. Nelson, J. F., Felicio, L. S., Randall, P. K., Sims, C. & Finch, C. E. A longitudinal study of estrous cyclicity in aging C57BL/6J mice: I. Cycle frequency, length and vaginal cytology. *Biol. Reprod.* **27**, 327–339 (1982).



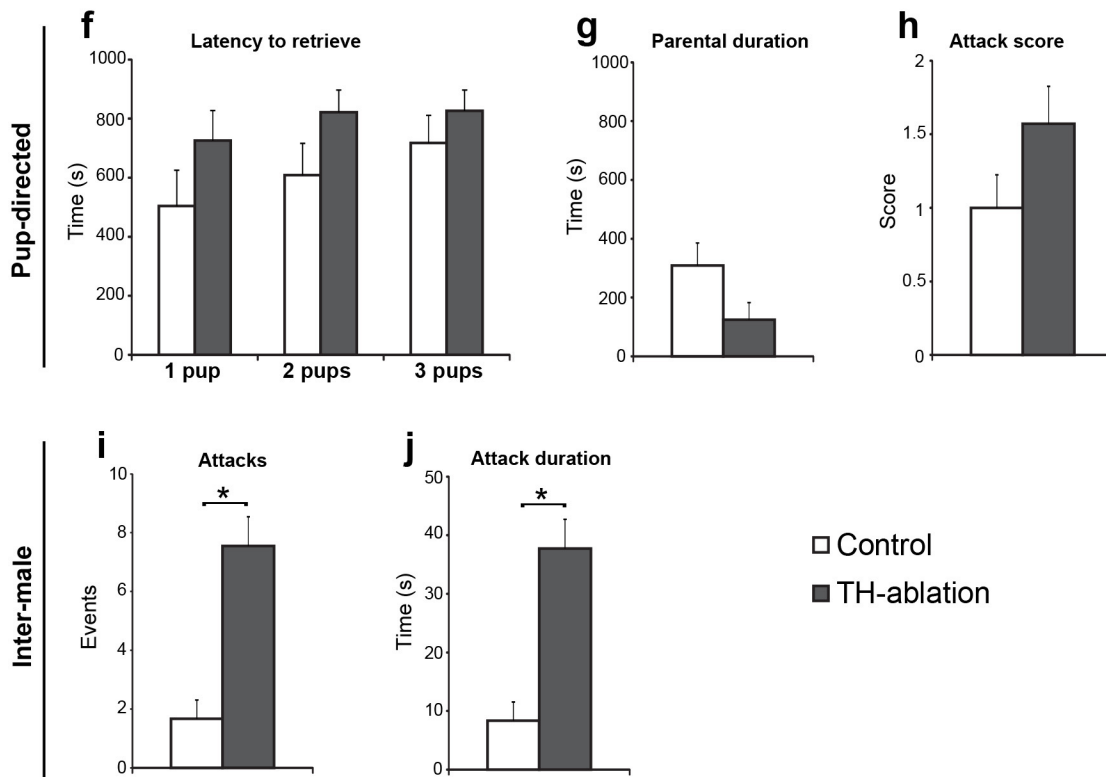
Extended Data Figure 1 | 6-OHDA injection into AVPV of male and female mice results in specific ablation of TH⁺ neurons. **a**, TH immunostaining in AVPVs of females and males injected with 6-OHDA (TH-ablation) or saline (control). Scale bars, 20 μ m. **b**, Number of TH⁺ AVPV neurons in TH-ablation and control females and males (females, $n_{\text{TH-ablation}} = 12$, $n_{\text{control}} = 13$; males, $n_{\text{ablation}} = 9$, $n_{\text{control}} = 8$; *** $P < 0.001$, two-way ANOVA with Fisher's

multiple comparisons). **c**, TH immunostaining in brain slices from females injected with 6-OHDA or saline into the AVPV. Scale bars, 20 μ m. **d**, Number of TH-immunoreactive neurons in TH-expressing brain areas ($n_{\text{ablation}} = 5$, $n_{\text{control}} = 5$). Data are means \pm s.e.m. AVPV, anteroventral periventricular nucleus, ARC, arcuate nucleus, SN, substantia nigra, VTA, ventral tegmental area.

Females



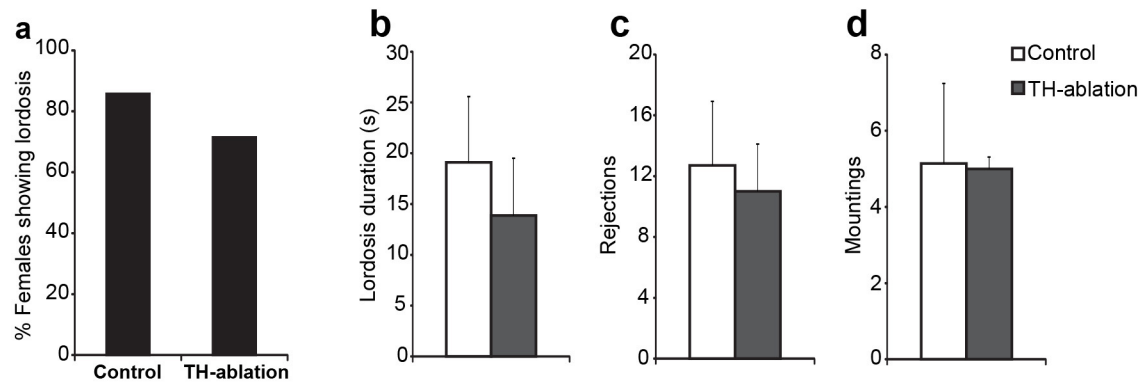
Males



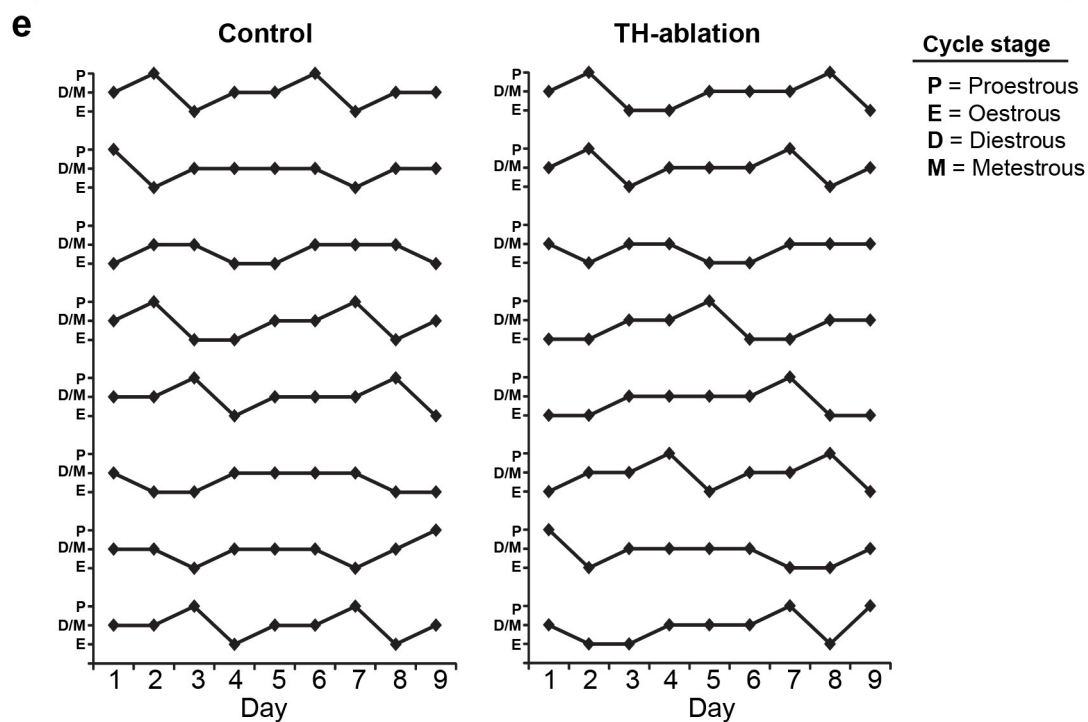
Extended Data Figure 2 | TH⁺ ablation in AVPV impairs maternal behaviour and increases inter-male aggression. a–d, Maternal behaviour of virgin females in TH-ablation and control groups ($n_{\text{ablation}} = 13$, $n_{\text{control}} = 13$). e, Pup retrieval of postpartum females in TH-ablation and control groups

($n_{\text{ablation}} = 11$, $n_{\text{control}} = 11$). f–h, Pup-directed behaviours of virgin (g, h) and parental (f) males in TH-ablation and control groups. i, j, Inter-male aggression in TH-ablated and control groups ($n_{\text{ablation}} = 10$, $n_{\text{control}} = 10$). Data are means + s.e.m. * $P < 0.05$; ** $P < 0.01$, Mann–Whitney U -test.

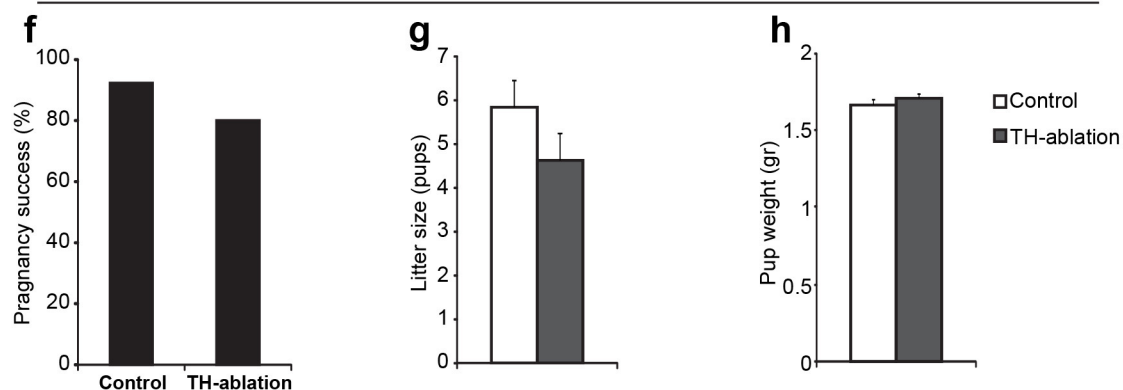
Female sexual behaviour



Oestrous cycle



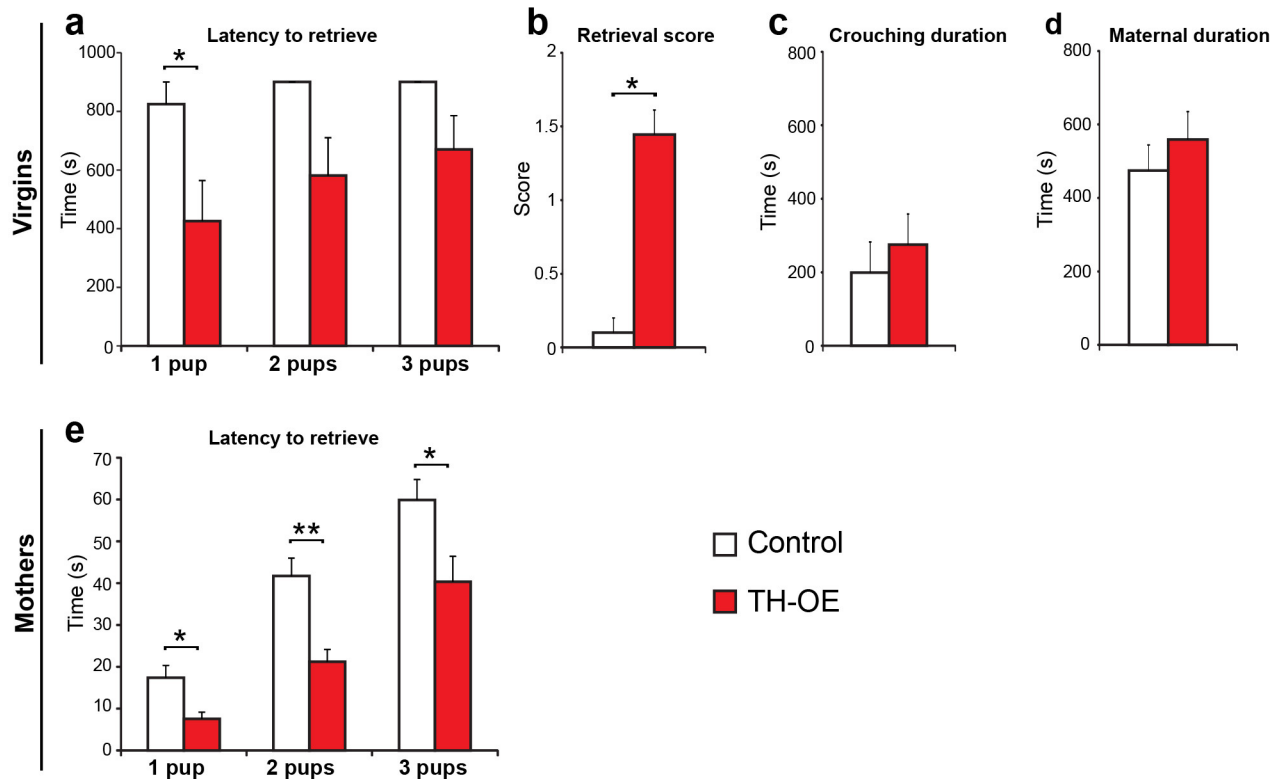
Female reproductive success



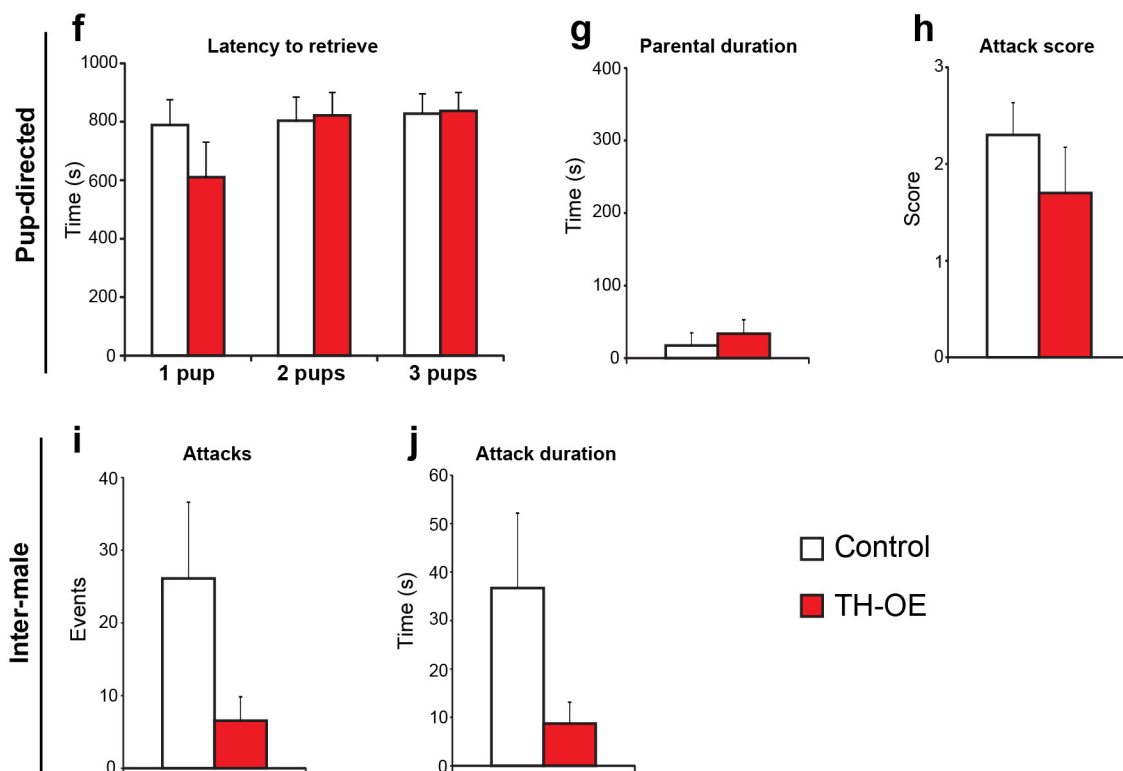
Extended Data Figure 3 | TH⁺ ablation in female AVPV does not affect sexual behaviour and reproduction. **a–d**, Female sexual behaviour in TH-ablation (6-OHDA) and control (saline) groups ($n_{\text{ablation}} = 7$, $n_{\text{control}} = 7$). **a**, Percentage of females displaying lordosis behaviour. **b**, Total duration of lordosis behaviour. **c**, Number of defensive rejections of the intruder male by the subject females. **d**, Successful sexual mounting events of the intruder

male on the subject females. **e**, Oestrus cycles in TH-ablation and control females ($n_{\text{ablation}} = 8$, $n_{\text{control}} = 8$). **f–h**, Female reproductive success in TH-ablation and control groups. **f**, Gestational success in percentage after copulation with males ($n_{\text{ablation}} = 14$, $n_{\text{control}} = 14$). Litter size (**g**) and pup weight (**h**) of TH-ablation and control females ($n_{\text{ablation}} = 12$, $n_{\text{control}} = 12$). Data are means + s.e.m.

Females

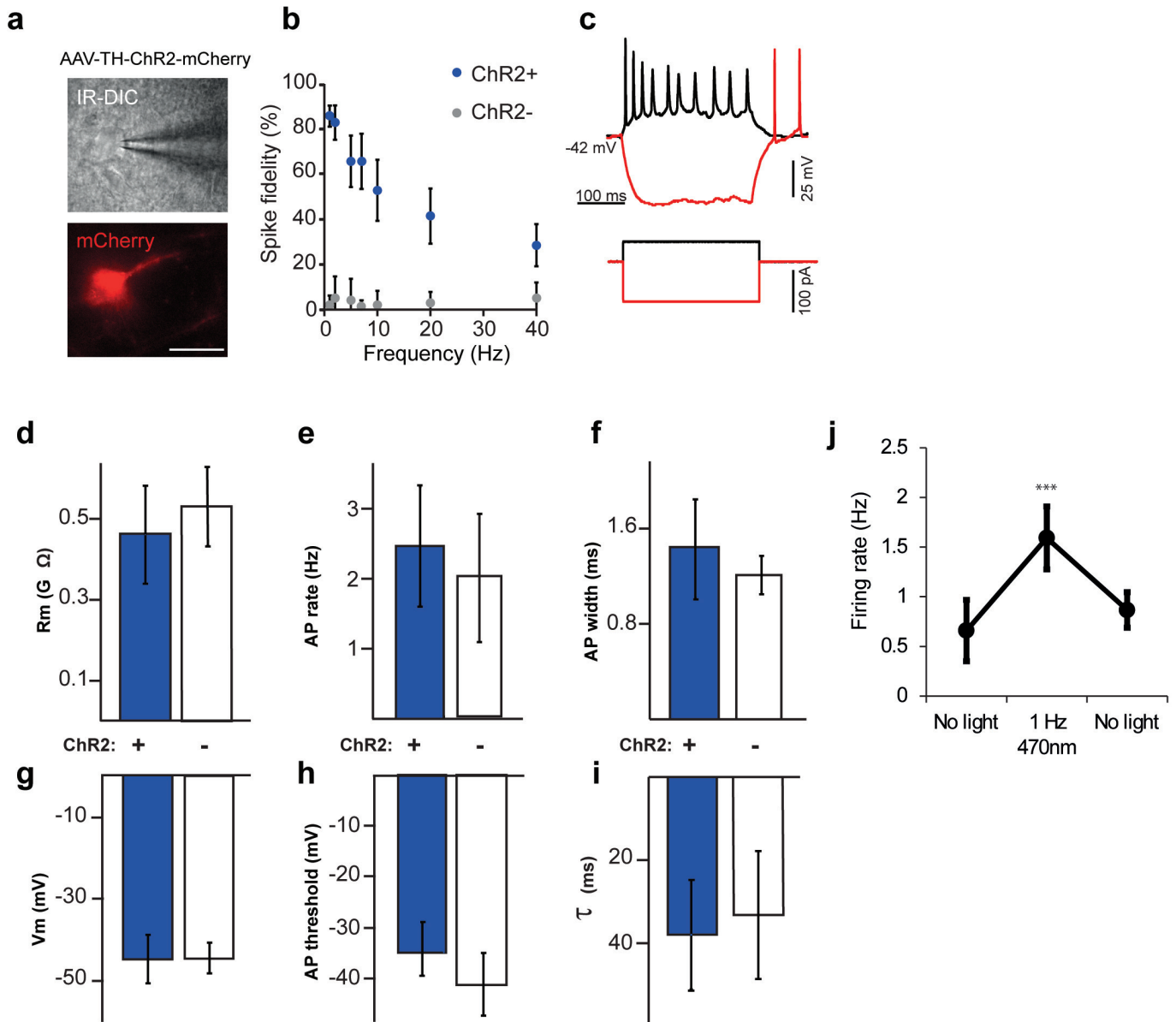


Males



Extended Data Figure 4 | TH overexpression in TH⁺ AVPV neurons increases maternal pup retrieval. **a–d**, Maternal behaviour in TH-overexpression (TH-OE) and control virgin females ($n_{OE} = 9$, $n_{control} = 10$). **e**, Pup retrieval in TH-OE and control postpartum females ($n_{OE} = 8$,

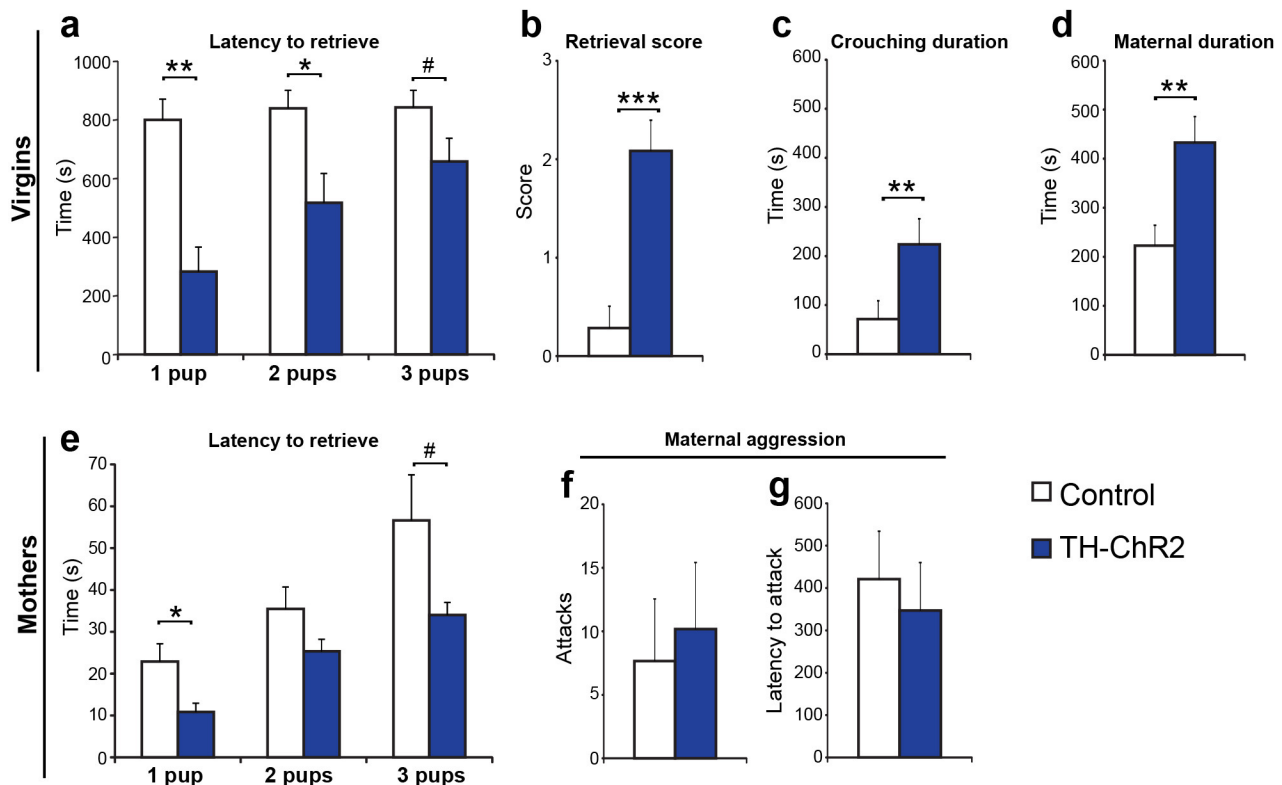
$n_{control} = 7$). **f–h**, Pup-directed behaviours in TH-OE and control virgin (**g, h**) and paternal (**f**) males. **i, j**, Inter-male aggression in TH-OE and control males ($n_{OE} = 10$, $n_{control} = 10$). Data are mean + s.e.m. * $P < 0.05$; ** $P < 0.01$, Mann–Whitney U -test.



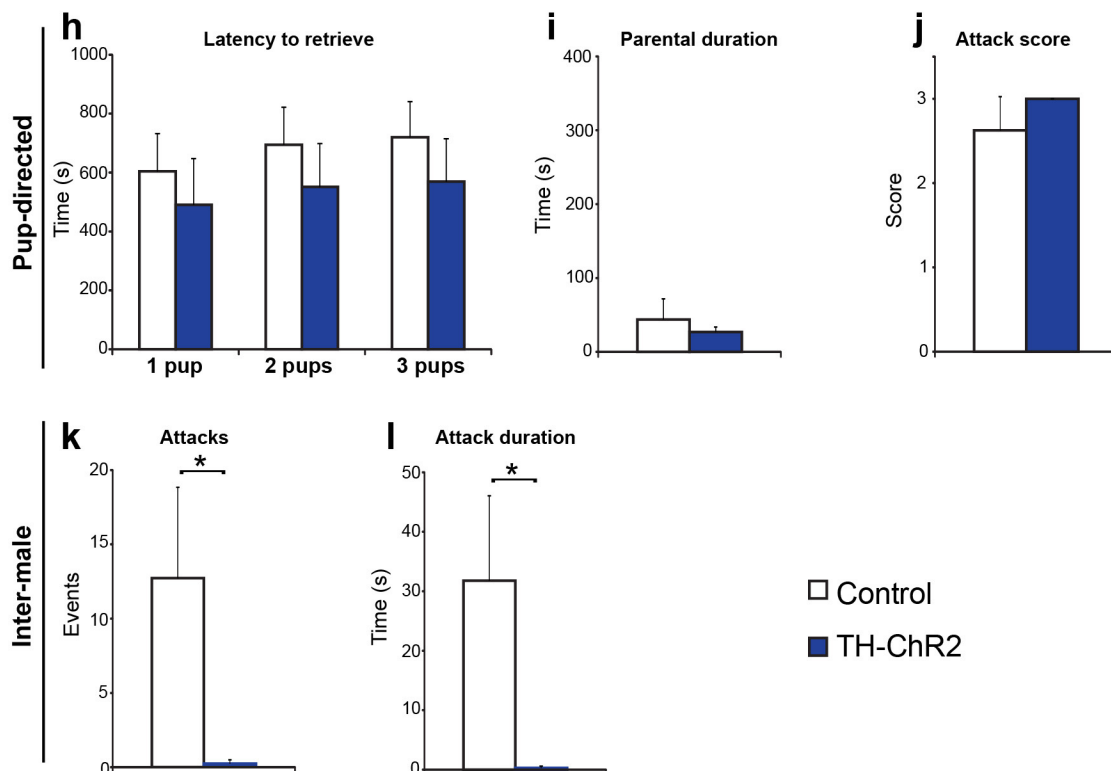
Extended Data Figure 5 | Intrinsic electrophysiological properties of TH⁺ AVPV neurons. Whole-cell recordings were performed in acute coronal slices from TH-Cre mice co-injected with DIO-EYFP and DIO-ChR2(E123T/T159C)-mCherry viral vectors. Cells were identified based on EYFP expression and recorded in current-clamp mode. Cells were then classified as ChR2⁺ or ChR2⁻ based on the presence or absence of a direct, short-latency (<1 ms) light-evoked photocurrent response. **a**, Differential interference contrast (top) and mCherry fluorescence (bottom) images of a TH⁺ AVPV cell expressing ChR2-mCherry. Scale bar, 20 μ m. **b**, Light-evoked spiking fidelity in TH⁺ AVPV neurons across varying light pulse frequencies (ChR2⁺, $n = 12$ cells; ChR2⁻, $n = 10$ cells). Light pulse trains containing 20 pulses (10 ms, 19 mW mm⁻², 475 nm) at each frequency were used to calculate response rates. Only spikes that occurred within 10 ms of light onset were calculated

as direct responses. Apparent responses in ChR2⁻ cells are attributed to the ongoing spontaneous firing of these neurons. **c**, Current clamp recording of voltage responses to negative (100 pA, red) and positive (50 pA, black) current injections in an AVPV TH⁺ neuron. **d–i**, Intrinsic electrical properties of TH⁺/ChR2⁺ (blue bars) and TH⁺/ChR2⁻ (white bars) cells, calculated from responses to current injections as shown in **c**: input resistance (**d**), spontaneous action potential firing rate (**e**), width of action potentials at half-maximum (**f**), resting membrane potential (**g**), action potential threshold (**h**) and membrane time constant (**i**). All showed no marked difference between ChR2⁺ and ChR2⁻ cells (ChR2⁺, $n = 8$; ChR2⁻, $n = 4$). **j**, Action potential firing rates of TH⁺/ChR2⁺ cells recorded in whole-cell patch clamp mode before, during and after 1 Hz optogenetic stimulation (data are means \pm s.e.m., *** $P < 0.05$, paired t -test, $n = 7$ cells).

Females



Males



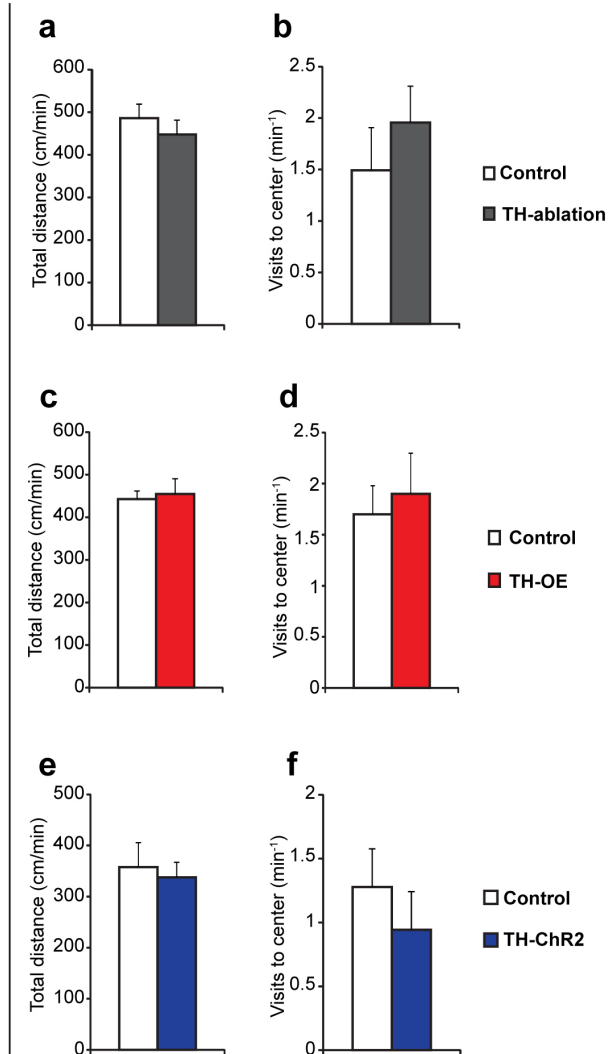
Extended Data Figure 6 | Optogenetic activation of TH⁺ AVPV neurons increases maternal behaviour and reduces inter-male aggression.

a–d, Maternal behaviour of virgin females during optogenetic activation in TH⁺ AVPV neurons ($n_{\text{ChR2}} = 12$, $n_{\text{control}} = 14$). **e–g**, Pup retrieval and maternal aggression of postpartum females during optogenetic activation in TH⁺ AVPV

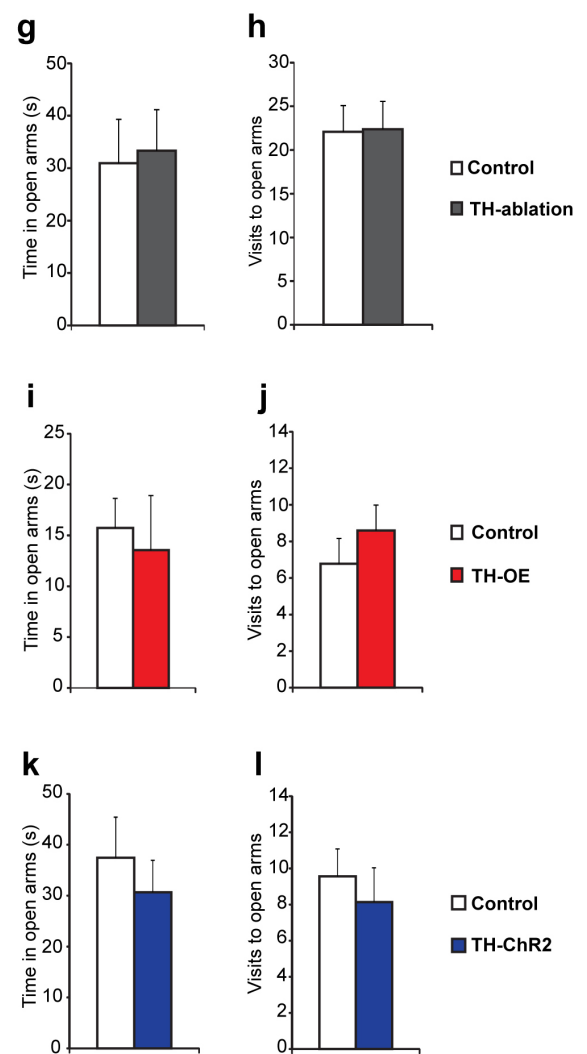
neurons ($n_{\text{ChR2}} = 6$, $n_{\text{control}} = 6$). **h–j**, Pup-directed behaviours of virgin (**i, j**) and paternal (**h**) males through optogenetic activation in TH⁺ AVPV neurons.

k, l, Inter-male aggression through optogenetic activation in TH⁺ AVPV neurons ($n_{\text{ChR2}} = 10$, $n_{\text{control}} = 10$). Data are means + s.e.m. # $P = 0.09$ (**a**) or 0.05 (**e**), * $P < 0.05$; ** $P < 0.01$, *** $P < 0.001$, Mann–Whitney *U*-test.

Open field assay



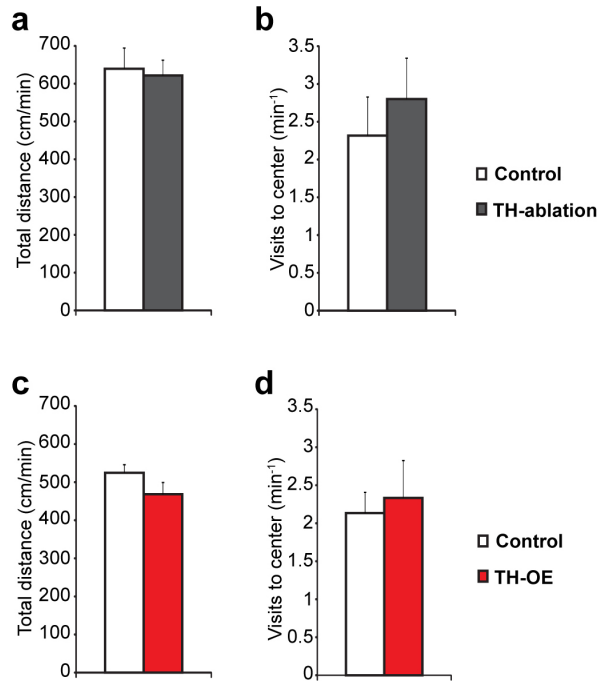
Elevated plus maze



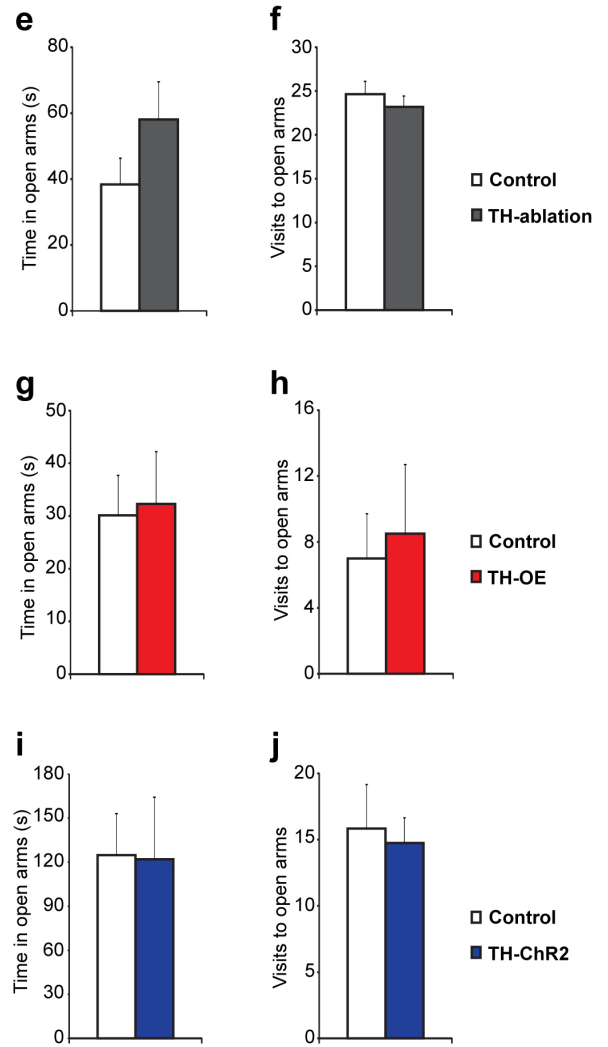
Extended Data Figure 7 | TH⁺ AVPV neuronal manipulations do not affect locomotion or anxiety in females. a–f, Open field assay. Total distance travelled (left) and total visits to centre of the field (right) for TH-ablation (top), TH-OE (middle) and TH-ChR2 (bottom) females relative to respective control groups. g–l, Elevated plus maze assay. Total time spent in the open arms of the maze (left) and total visits to open arms (right) for TH-ablation (top),

TH-OE (middle) and TH-ChR2 (bottom) females relative to respective control groups (TH-ablation, $n_{\text{ablation}} = 13$, $n_{\text{control}} = 13$; TH-OE, $n_{\text{OE}} = 10$, $n_{\text{control}} = 10$; TH-ChR2, $n_{\text{ChR2}} = 7$, $n_{\text{control}} = 9$). TH-ChR2 mice and EYFP-expressing controls were tested using attached fibre optics and blue light stimulation as described for maternal behaviour testing. Data are means + s.e.m.

Open field assay

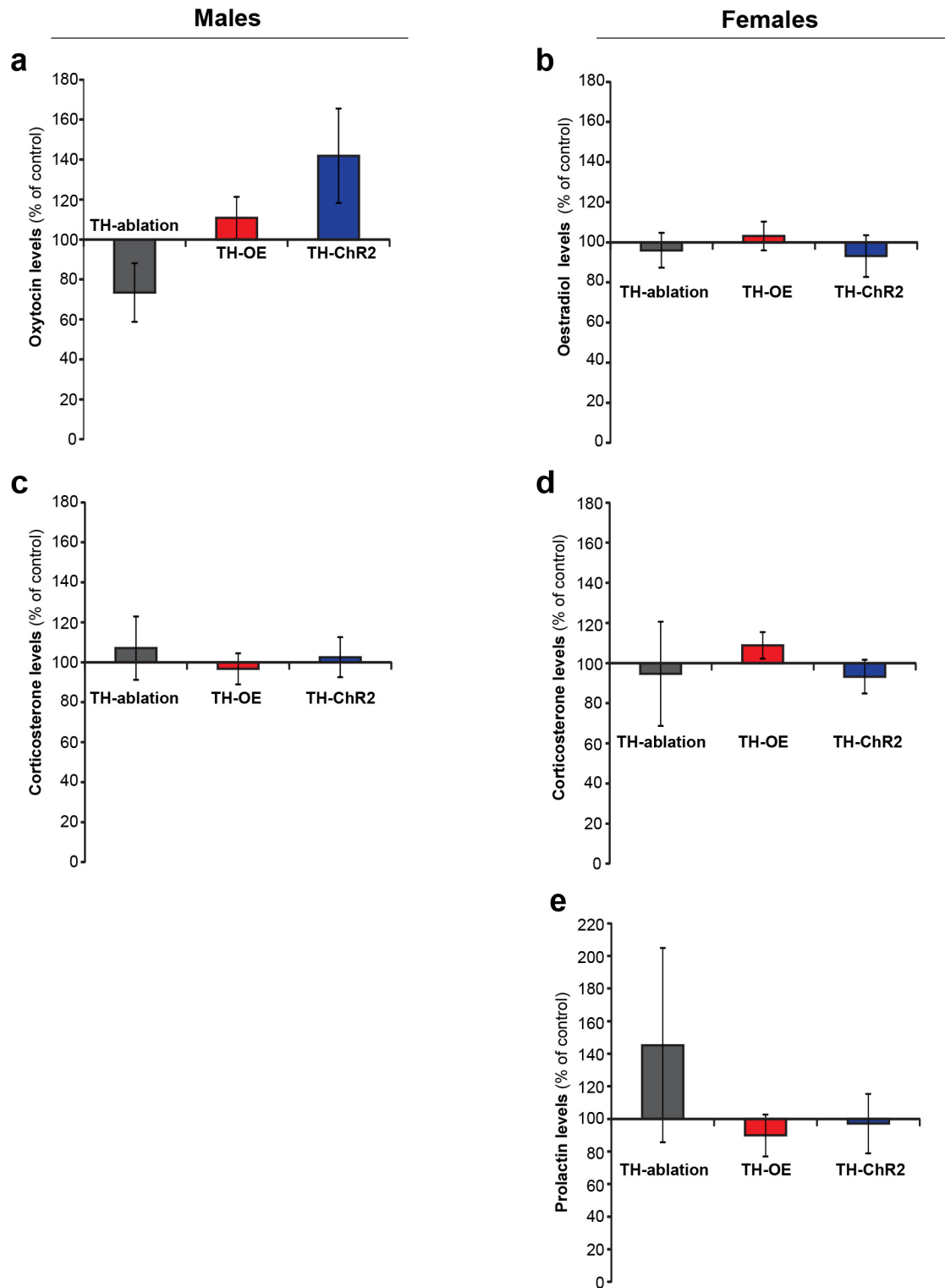


Elevated plus maze



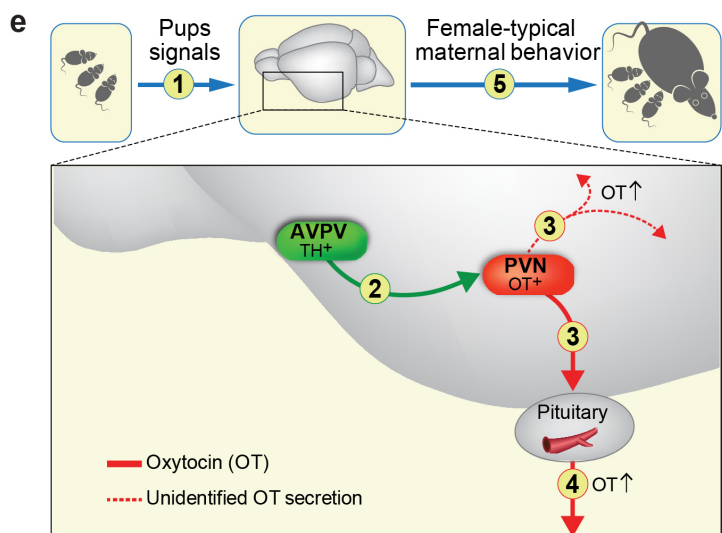
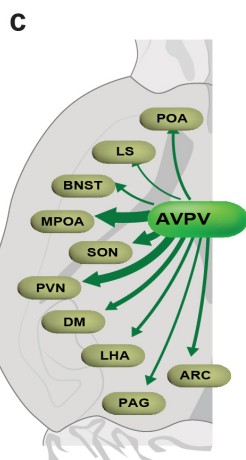
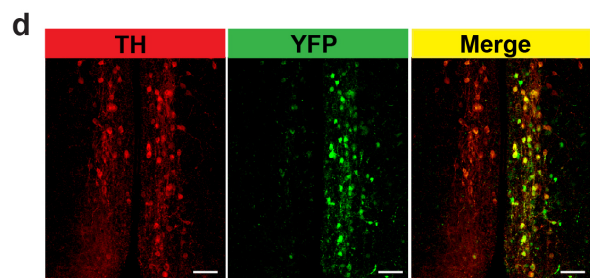
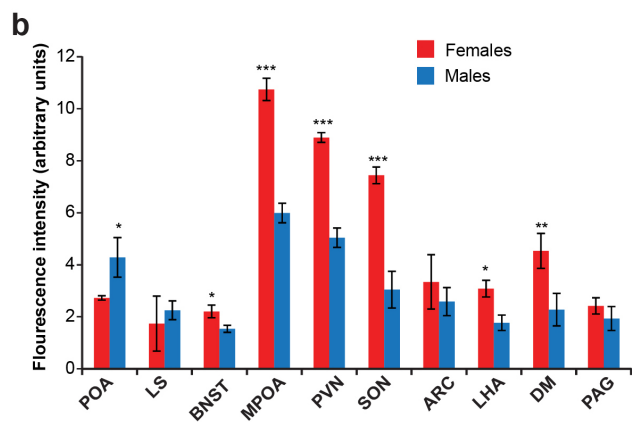
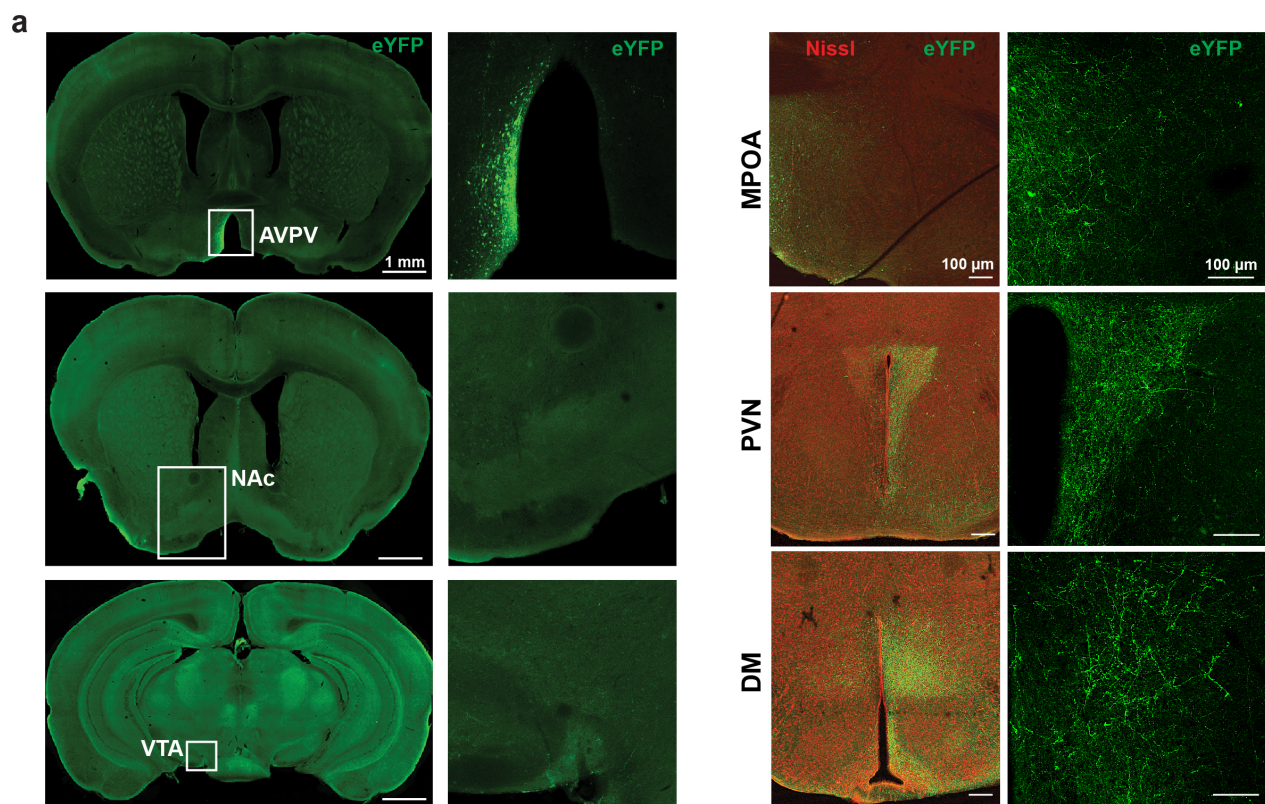
Extended Data Figure 8 | TH⁺ AVPV neuronal manipulations do not affect locomotion or anxiety in males. **a–d**, Open field assay. Total distance travelled (left) and total visits to centre of the field (right) for TH-ablation (top) and TH-OE (bottom) males relative to control groups. **e–j**, Elevated plus maze assay. Total time spent in open arms of the maze (left) and total visits to open arms (right) for TH-ablation (top), TH-OE (middle) and ChR2 (bottom)

males relative to respective control groups. (TH-ablation, $n_{ablation} = 12$, $n_{control} = 12$; TH-OE, $n_{OE} = 10$, $n_{control} = 10$; TH-ChR2, $n_{ChR2} = 4$, $n_{control} = 7$). TH-ChR2 mice and EYFP-expressing controls were tested with attached fibre optics and blue light stimulation as described for parental behaviour testing. Data are means + s.e.m.



Extended Data Figure 9 | Hormones levels in plasma of TH⁺ AVPV manipulated mice. **a**, OT levels in males (TH-ablation, $n_{\text{ablation}} = 9$, $n_{\text{control}} = 9$; TH-OE, $n_{\text{OE}} = 6$, $n_{\text{control}} = 5$; TH-ChR2, $n_{\text{ChR2}} = 7$, $n_{\text{control}} = 8$). **b**, Oestradiol levels in females (TH-ablation, $n_{\text{ablation}} = 9$, $n_{\text{control}} = 11$; TH-OE, $n_{\text{OE}} = 6$, $n_{\text{control}} = 5$; TH-ChR2, $n_{\text{ChR2}} = 10$, $n_{\text{control}} = 10$). **c**, **d**, Corticosterone levels in females (**c**) and males (**d**) (females: TH-ablation, $n_{\text{ablation}} = 6$, $n_{\text{control}} = 5$; TH-OE, $n_{\text{OE}} = 9$, $n_{\text{control}} = 9$; TH-ChR2, $n_{\text{ChR2}} = 9$,

$n_{\text{control}} = 8$; males: TH-ablation, $n_{\text{ablation}} = 4$, $n_{\text{control}} = 4$; TH-OE, $n_{\text{OE}} = 8$, $n_{\text{control}} = 8$; TH-ChR2, $n_{\text{ChR2}} = 7$, $n_{\text{control}} = 7$). **e**, Prolactin levels in females (TH-ablation, $n_{\text{ablation}} = 11$, $n_{\text{control}} = 10$; TH-OE, $n_{\text{OE}} = 5$, $n_{\text{control}} = 5$; TH-ChR2, $n_{\text{ChR2}} = 6$, $n_{\text{control}} = 6$). Data are normalized to matched control groups. No significant differences were found between the TH-manipulated and control groups in the presented parameters. Data are means \pm s.e.m.



Extended Data Figure 10 | TH⁺ AVPV neuronal projection and suggested model by which TH⁺ AVPV neurons promote maternal care. **a**, Coronal brain sections of mice unilaterally injected with a conditional EYFP-expressing viral vector (AAV-DIO-EYFP) into the AVPV of TH-Cre female mice. Projections from TH⁺ AVPV neurons into various brain structures are presented. Scale bars, 1 mm (left panel) and 100 μ m (right panel). **b**, Fluorescent intensities of EYFP-labelled projection fibres of TH⁺ AVPV neurons in different brain structures of TH-Cre females and males injected with a conditional EYFP-expressing viral vector (data are means \pm s.e.m., females, $n = 3$; males, $n = 5$, $*P < 0.05$, $**P < 0.01$, $***P < 0.001$, two-tailed Student's t -test). AVPV, anteroventral periventricular nucleus; POA, preoptic area; LS, lateral septum; BNST, bed nucleus of the stria terminalis; MPOA, medial preoptic area; SON, supraoptic nucleus; PVN, paraventricular nucleus;

DM, dorsomedial nucleus; LHA, lateral hypothalamic area; PAG, periaqueductal grey; ARC, arcuate nucleus. **c**, Schematic illustration of projections from TH⁺ AVPV neurons of adult females, in a transverse view. Arrow thickness indicates projection density, measured as fluorescent intensity of fibres labelled with EYFP. **d**, Specificity of Cre-dependent EYFP expression. Image showing a coronal section from a TH-Cre mouse injected with AAV-DIO-EYFP into the AVPV. Images show colocalization of EYFP in TH⁺ neurons (green) and immunostaining for TH (red). Scale bars, 50 μ m. **e**, Suggested model by which TH⁺ AVPV neurons promote female-typical OT release and maternal behaviour. (1) Pup-related sensory signals induce changes in the activity of TH⁺ AVPV neurons; (2) activated TH⁺ AVPV neurons stimulate OT⁺ PVN neurons; (3, 4) OT⁺ PVN neurons secrete OT into central nervous system and blood; (5) maternal behaviour is facilitated.

Cell-fate determination by ubiquitin-dependent regulation of translation

Achim Werner^{1,2}, Shintaro Iwasaki², Colleen A. McGourty², Sofia Medina-Ruiz², Nia Teerikorpi^{1,2}, Indro Fedrigo², Nicholas T. Ingolia² & Michael Rape^{1,2}

Metazoan development depends on the accurate execution of differentiation programs that allow pluripotent stem cells to adopt specific fates¹. Differentiation requires changes to chromatin architecture and transcriptional networks, yet whether other regulatory events support cell-fate determination is less well understood. Here we identify the ubiquitin ligase CUL3 in complex with its vertebrate-specific substrate adaptor KBTBD8 (CUL3^{KBTBD8}) as an essential regulator of human and *Xenopus tropicalis* neural crest specification. CUL3^{KBTBD8} monoubiquitylates NOLC1 and its paralogue TCOF1, the mutation of which underlies the neurocristopathy Treacher Collins syndrome^{2,3}. Ubiquitylation drives formation of a TCOF1–NOLC1 platform that connects RNA polymerase I with ribosome modification enzymes and remodels the translational program of differentiating cells in favour of neural crest specification. We conclude that ubiquitin-dependent regulation of translation is an important feature of cell-fate determination.

Cullin-RING ligases (CRLs), the largest class of ubiquitylation enzymes, have critical roles in metazoan development^{4–10}. CRLs recognize their substrates through ~300 adaptor proteins, several of which are differentially expressed during development^{11–13}. Although mutations in CRL adaptors have been linked to human pathology^{14–16}, little

is known about how distinct CRLs ensure robust differentiation into specialized cell types.

To discover CRLs with crucial roles in development, we employed genome-wide transcript analysis of differentiating human embryonic stem cells (hESCs). These experiments revealed a strong reduction in the abundance of the vertebrate-specific CUL3 adaptor KBTBD8 during hESC differentiation (Extended Data Fig. 1a–c), which we confirmed for KBTBD8 messenger RNA and protein by quantitative reverse transcription PCR (qRT-PCR) and western blot analysis (Extended Data Fig. 1d–g). Consistent with evolutionary conservation, downregulation of KBTBD8 was observed in differentiating mouse ESCs (Extended Data Fig. 1h, i), as well as during *Xenopus tropicalis* development¹⁷.

Depletion of KBTBD8 did not affect the cell cycle, survival, or pluripotency programs of hESCs (Extended Data Fig. 2a–e). Instead, gene expression profiles of hESCs subjected to embryoid body differentiation suggested that KBTBD8 was required for neural crest specification (Extended Data Fig. 2f and Supplementary Table 1). qRT-PCR experiments confirmed that loss of KBTBD8 reduced expression of neural crest markers, including *FOXD3* and *SOX10*, which was accompanied by an increase in transcripts associated with

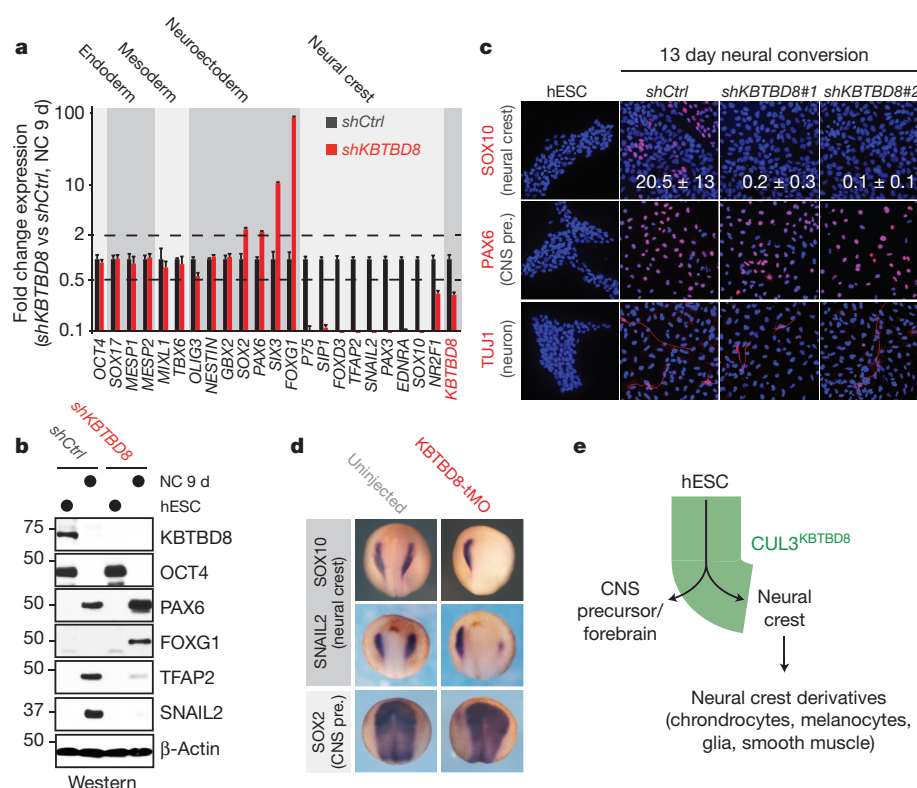


Figure 1 | CUL3^{KBTBD8} drives neural crest specification. **a**, hESCs stably depleted of KBTBD8 were subjected to neural conversion and analysed by qRT-PCR (mean of 3 technical replicates, \pm s.e.m.). **b**, Depletion of KBTBD8 results in loss of neural crest cells, as determined by western blot analysis (full scans available in Supplementary Fig. 1). NC 9d, neural conversion, day 9; molecular weight is given in kDa. **c**, KBTBD8-depleted hESCs were subjected to neural conversion and analysed by immunofluorescence microscopy (mean of 3 biological replicates, \pm s.e.m.; ~1,500 cells per condition). **d**, *Xenopus tropicalis* embryos injected with translation-blocking morpholinos against KBTBD8 were analysed by *in situ* hybridization. **e**, Model of the CUL3^{KBTBD8}-controlled developmental switch.

¹Howard Hughes Medical Institute, University of California, Berkeley, California 94720, USA. ²Department of Molecular and Cell Biology, University of California, Berkeley, California 94720, USA.

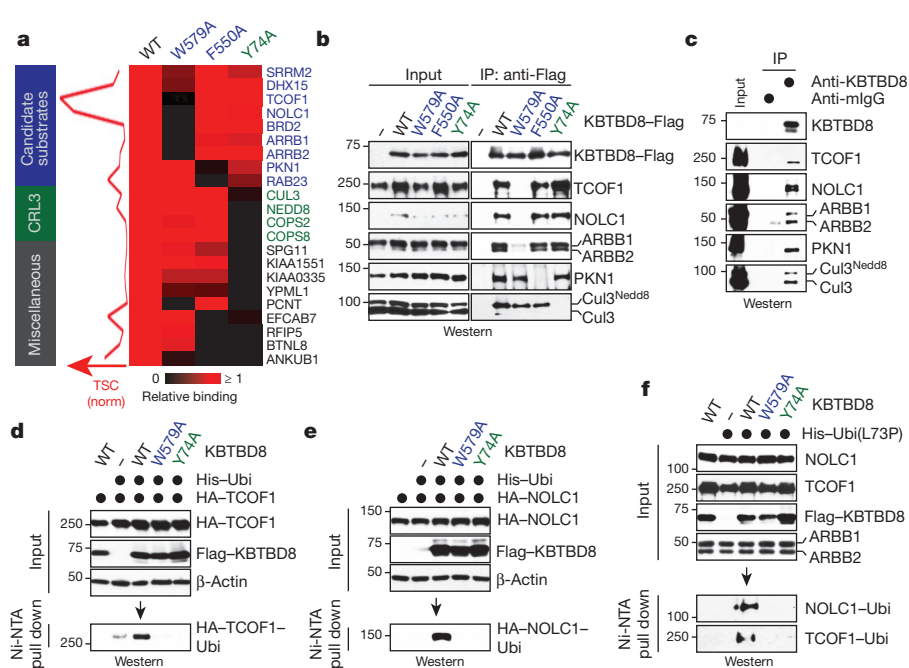


Figure 2 | CUL3^{KBTBD8} monoubiquitylates TCOF1 and NOLC1. **a**, High-confidence interactors of wild-type (WT) or mutant KBTBD8. Left: normalized total spectral counts (TSCs) per interactor of wild-type KBTBD8 (sum of 3 biological replicates per condition). Right: heat map depicting binding relative to wild-type KBTBD8. **b**, Verification of KBTBD8 interactions in 293T cells by anti-Flag immunoprecipitation and western blot analysis. IP, immunoprecipitation; molecular weight is given in kDa. **c**, Immunoprecipitation of KBTBD8 from hESCs (full scans available in Supplementary Fig. 1). **d**, Ubiquitylated (Ubi) HA-tagged TCOF1 detected after denaturing Ni-NTA purification in 293T cells reconstituted with KBTBD8 variants. **e**, Monoubiquitylation of HA-NOLC1 by CUL3^{KBTBD8} in 293T cells. **f**, Monoubiquitylation of endogenous TCOF1 and NOLC1 in 293T cells reconstituted with KBTBD8 variants and His-ubiquitin(L73P).

central nervous system (CNS) precursor and forebrain identity (FOXG1, SIX3; Extended Data Fig. 2g).

On the basis of these observations, we subjected hESCs to dual-SMAD inhibition ('neural conversion'), which directs differentiation towards CNS precursor and neural crest cells¹⁸. As seen during embryoid body differentiation, depletion of KBTBD8 caused a striking loss of neural crest cells and an increase in CNS precursors (Fig. 1a, b), which was seen for multiple short hairpin RNAs (shRNAs) and was rescued by shRNA-resistant KBTBD8 (Fig. 3b and Extended Data Fig. 3g). We corroborated these results with single-cell resolution using the neural crest marker SOX10 (Fig. 1c) or AP2, p75 and HNK1, which are co-expressed in most neural crest cells (Extended Data Fig. 3a). KBTBD8 was required for early neural crest specification, with CNS precursor markers accumulating in KBTBD8-depleted cells when neural crest markers were first detected in control experiments (Extended Data Fig. 3b–h). KBTBD8 was accordingly critical for differentiation of hESC-derived neural crest cells into glia, mesenchymal cells, melanocytes, or chondrocytes (Extended Data Fig. 4a, b). Also in *Xenopus tropicalis*, downregulation or inhibition of CUL3^{KBTBD8} prevented neural crest formation and caused an expansion of the CNS precursor territory in the manipulated part of the embryo (Fig. 1d and Extended Data Fig. 4c). Thus, CUL3^{KBTBD8} regulates a developmental switch that controls the generation of the neural crest, an embryonic cell population that is found only in vertebrates (Fig. 1e).

To isolate essential targets of CUL3^{KBTBD8}, we used CompPASS mass spectrometry to capture proteins that bound wild-type KBTBD8 but not variants with a mutant substrate-binding domain (KBTBD8(W579A); Extended Data Fig. 5a–d). These interaction networks identified the paralogs NOLC1 and TCOF1 as predominant interactors of KBTBD8, which were not recognized by KBTBD8(W579A) (Fig. 2a). Using western blot analysis, we confirmed binding of TCOF1 and NOLC1 to KBTBD8 but not KBTBD8(W579A) (Fig. 2b), and showed that the same association occurred between endogenous proteins in hESCs (Fig. 2c) and in reconstituted *in vitro* systems (Extended Data Fig. 5e, f). Denaturing purification of ubiquitin conjugates revealed that KBTBD8, but neither KBTBD8(W579A) nor CUL3-binding-deficient KBTBD8(Y74A), induced the robust monoubiquitylation of TCOF1 and NOLC1 (Fig. 2d–f). These events required a cofactor, β -arrestin, the depletion of which prevented KBTBD8 recognition and monoubiquitylation of TCOF1 and NOLC1 (Extended Data Fig. 5g–j).

Similar to loss of KBTBD8, hESCs expressing only KBTBD8(W579A) or KBTBD8(Y74A) failed to support neural crest specification and showed increased abundance of CNS precursors

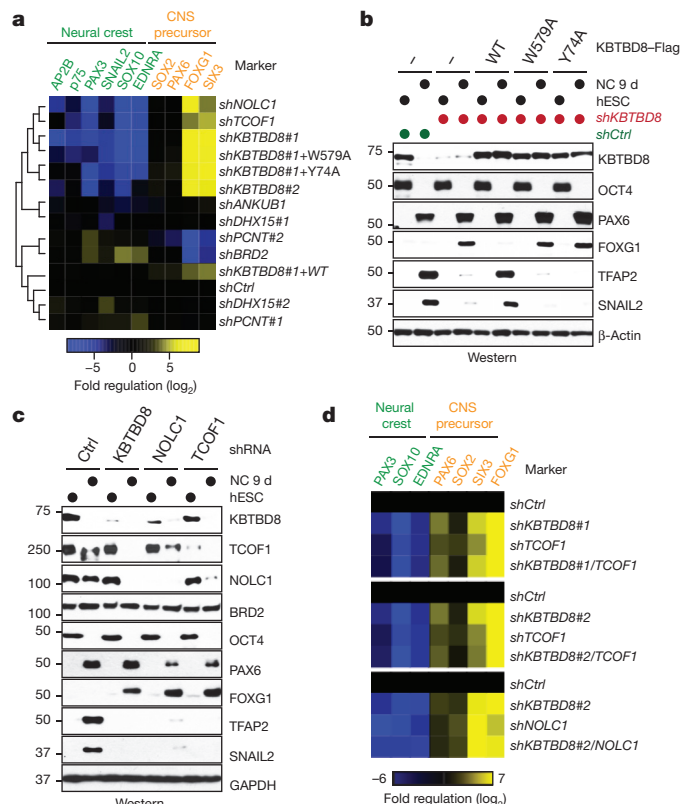


Figure 3 | CUL3^{KBTBD8} controls neural crest specification through TCOF1 and NOLC1. **a**, hESCs were reconstituted with shRNA-resistant KBTBD8 variants or depleted of KBTBD8-binding partners, subjected to neural conversion (9 d), and analysed by qRT-PCR and unsupervised clustering. **b**, Protein expression during neural conversion of hESCs reconstituted with shRNA-resistant KBTBD8 variants (full scans available in Supplementary Fig. 1). Molecular weight is shown in kDa. **c**, Protein expression in hESCs stably depleted of KBTBD8, TCOF1, or NOLC1 and subjected to neural conversion. **d**, hESCs were stably depleted of the indicated combinations of KBTBD8, TCOF1, or NOLC1, subjected to neural conversion (9 days), and analysed by qRT-PCR.

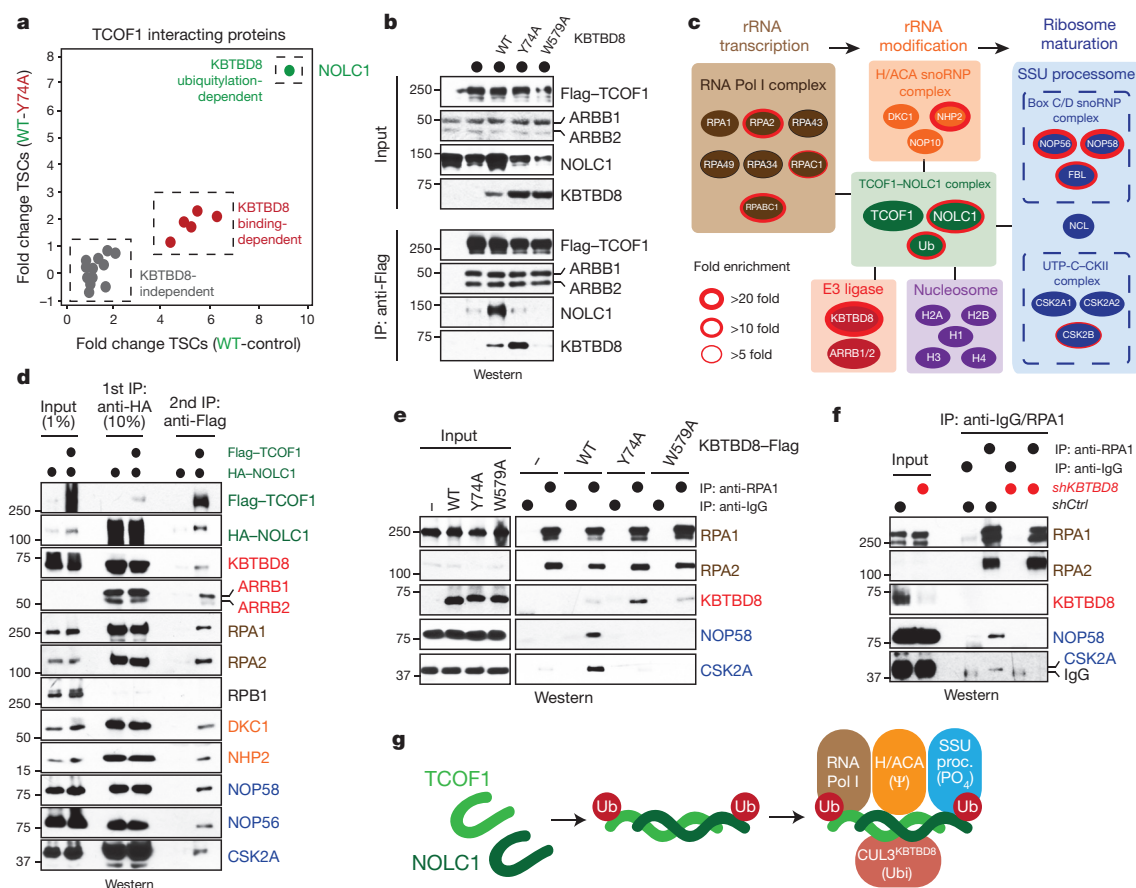


Figure 4 | Ubiquitylation-dependent TCOF1-NOLC1 complexes couple RNA polymerase I to ribosome modification enzymes. **a**, Interactors of TCOF1 in 293T cells reconstituted with KBTBD8 or KBTBD8(Y74A) (sum of 3 biological replicates per condition). **b**, Validation of CUL3^{KBTBD8}-dependent formation of TCOF1-NOLC1 complexes. **c**, CompPASS mass spectrometry analysis of sequential immunoprecipitation of Flag-TCOF1/HA-NOLC1 complexes. **d**, Validation of sequential affinity purification of KBTBD8-

dependent TCOF1-NOLC1 complexes (full scans available in Supplementary Fig. 1). **e**, Immunoprecipitation of RNA polymerase I from 293T cells reconstituted with KBTBD8 variants. **f**, Immunoprecipitation of RNA polymerase I from hESCs depleted of KBTBD8. **g**, Model of ubiquitin-dependent formation of a TCOF1-NOLC1 platform. Molecular weight is given in kDa.

(Fig. 3a, b and Extended Fig. 6a, b). The same aberrant differentiation program was observed if we depleted TCOF1 or NOLC1 (Fig. 3a, c and Extended Data Fig. 6a, c, d), but not other KBTBD8-binding partners (Fig. 3a and Extended Data Fig. 6e, f). Demonstrating that these proteins act in a common pathway, co-depletion of KBTBD8 and TCOF1 or NOLC1, respectively, mirrored the differentiation program of singly depleted hESCs (Fig. 3d). We therefore conclude that TCOF1 and NOLC1 are critical monoubiquitylation substrates of CUL3^{KBTBD8} during neural crest specification. Consistent with this notion, mutations in *TCOF1* cause Treacher Collins syndrome, a craniofacial disorder characterized by loss of cranial neural crest cells^{2,3}.

To understand how CUL3^{KBTBD8} drives neural crest specification, we identified proteins that selectively recognized ubiquitylated, but not unmodified, TCOF1 using cells that were reconstituted with either wild-type KBTBD8, inactive KBTBD8(Y74A), or empty vector. Notably, NOLC1 emerged as the major effector that was recruited to ubiquitylated TCOF1 (Fig. 4a), an observation that was confirmed by affinity purification coupled to western blot analysis (Fig. 4b). Monoubiquitylation often stabilizes binding partners, and depletion of KBTBD8 caused degradation of both TCOF1 and NOLC1 at later stages of neural conversion (Fig. 3c and Extended Data Fig. 7a, b).

On the basis of these results, we established a sequential affinity purification protocol to determine the composition of ubiquitylation-dependent TCOF1-NOLC1 complexes. We found that TCOF1-NOLC1 assemblies engaged RNA polymerase I; the H/ACA complex catalysing rRNA pseudouridylation; and the SSU processome controlling

maturation and modification of the small ribosomal subunit (Fig. 4c, d and Extended Data Fig. 7c). Accordingly, ubiquitylation by CUL3^{KBTBD8} brought endogenous RNA polymerase I into complexes with the SSU processome (Fig. 4e), which required TCOF1 and NOLC1 (Extended Data Fig. 7d). Similar observations were made in hESCs, where a robust interaction between RNA polymerase I and SSU processome was lost upon depletion of KBTBD8 (Fig. 4f). Thus, CUL3^{KBTBD8} induces the ubiquitin-dependent formation of TCOF1-NOLC1 complexes that serve as a platform to connect RNA polymerase I with enzymes responsible for ribosomal processing and modification (Fig. 4g). This observation supports a role of ubiquitylation in neural crest specification, as mutations in RNA polymerase I also cause Treacher Collins syndrome¹⁹.

Although KBTBD8 targets proteins linked to ribosome biogenesis, its depletion did not affect the abundance of rRNAs or mRNAs encoding ribosomal proteins; levels of ribosomal proteins; processing of precursor rRNAs; nucleolar integrity; export of the small ribosomal subunit; ribosome binding to mRNA judged by polysome gradient analysis; global mRNA translation detected by metabolic labelling; or cell survival (Fig. 5a, d and Extended Data Fig. 8a–h). Accordingly, a global reduction in translation caused by rapamycin did not phenocopy the loss of KBTBD8 (Extended Data Fig. 9a, b). Depletion of TCOF1 also did not affect rRNA synthesis, p53 activation, or cell survival at the time of neural crest specification (Extended Data Fig. 9c–e), although consistent with previous work²⁰, it reduced rRNA levels and triggered cell death at late stages of neural conversion

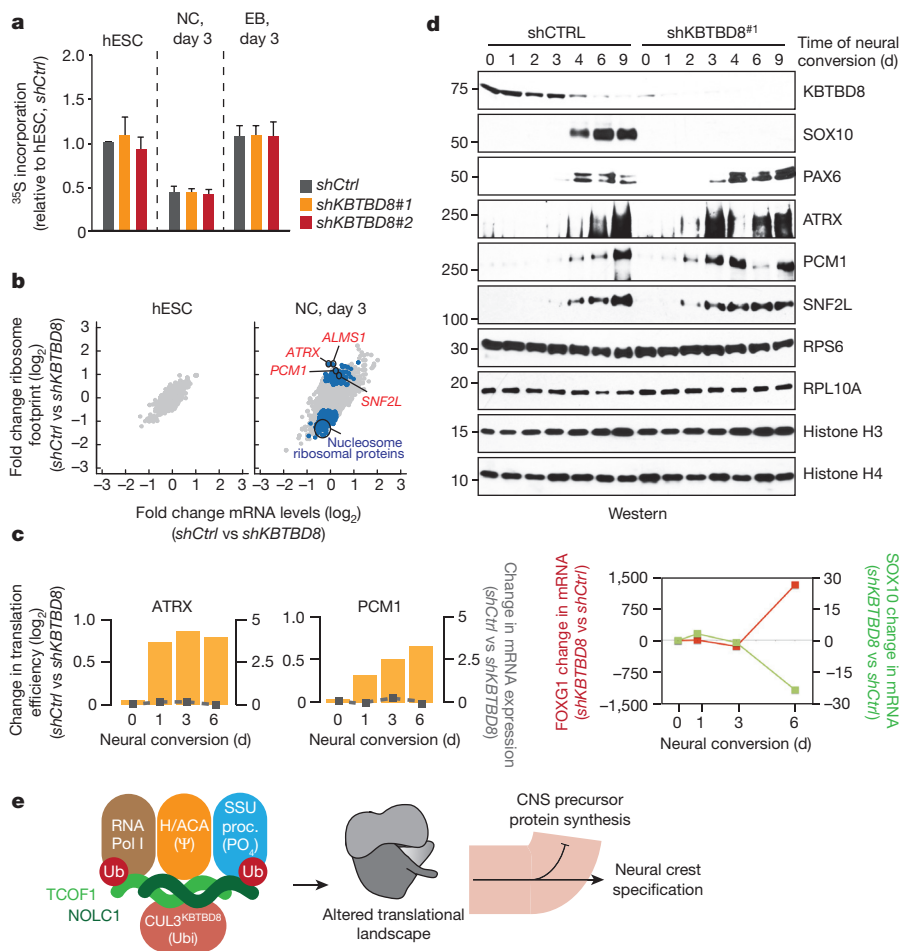


Figure 5 | CUL3^{KBTBD8} remodels translational programs during differentiation. **a**, ³⁵S-methionine incorporation into newly synthesized proteins in hESCs lacking KBTBD8 (3 biological replicates \pm s.e.m.). **b**, RNA-seq and ribosomal profiling of hESCs lacking KBTBD8 (blue, transcripts significantly changed in ribosome footprints but not mRNA; $n > 7,725$, $q < 0.1$). **c**, Translation efficiency of ATRX1 and PCM1 in hESCs lacking KBTBD8. Orange, translation efficiency change; grey inserts, mRNA abundance change (mean of two biological replicates). **d**, Protein expression in hESCs lacking KBTBD8 during neural conversion (full scans available in Supplementary Fig. 1; molecular weight is given in kDa). **e**, Model of CUL3^{KBTBD8}-dependent neural crest specification.

(Extended Data Fig. 9f, g). These observations implied that KBTBD8 and TCOF1 initially specify neural crest fate without altering ribosome abundance, global mRNA translation, or cell survival.

We therefore considered the possibility that CUL3^{KBTBD8}-dependent assembly of a ribosome modification platform might produce ribosomes with distinct translational output. Indeed, as seen by RNA sequencing and ribosome profiling, depletion of KBTBD8 changed the translational program of cells undergoing neural conversion, whereas it had no effect on protein synthesis in hESCs (Fig. 5b and Extended Data Fig. 10a). Similar observations were made for TCOF1, and the translation efficiency profiles of differentiating hESCs lacking KBTBD8 or TCOF1 were correlated (Extended Data Fig. 10b). Loss of KBTBD8 caused changes in translation immediately after initiation of differentiation, and thus, before specification of hESCs into neural crest or CNS precursor cells (Fig. 5c).

Analysis of regulated mRNAs showed that KBTBD8 suppressed the production of proteins specifying CNS precursors, whereas it did not affect translation of mRNAs connected to neural crest specification (Extended Data Fig. 10c). In this manner, KBTBD8 or TCOF1 delayed the accumulation of CNS precursor proteins, including ATRX and PCM1, until neural crest specification had occurred (Fig. 5d and Extended Data Fig. 10d). Underscoring the role of translational control, KBTBD8 enforced the correct timing of ATRX and PCM1 production without regulating their mRNA levels or protein stability (Fig. 5c, d and Extended Data Fig. 10d–f). The depletion of KBTBD8 also reduced the translation of mRNAs encoding histones or ribosomal components (Fig. 5b), yet as expected from their long half-lives^{21,22}, the levels of the corresponding proteins were not diminished during our differentiation experiment (Fig. 5d). Thus, the CUL3^{KBTBD8}-dependent formation of a ribosome modification platform alters translation of specific

mRNAs, which delays accumulation of CNS precursor proteins until hESCs have accomplished neural crest specification.

Our work documents an important role for ubiquitylation in remodelling translational programs during differentiation, and defines an early function for the CUL3^{KBTBD8} substrate and Treacher Collins syndrome-associated protein TCOF1 in neural crest specification (Fig. 5e). We hypothesize that CUL3^{KBTBD8} and TCOF1 may govern the production of differentially modified ribosomes, potentially including post-transcriptional changes in rRNA pseudouridylation and base methylation, or phosphorylation and ubiquitylation of ribosomal proteins or ribosome-associated factors. Such modifications may affect the interactions of ribosomes with select mRNAs, with factors that deliver specific mRNAs to the ribosome, or with proteins that control the synthesis or degradation of distinct mRNAs^{23–25}.

Together with studies implying specific functions for ribosomal proteins during differentiation^{26,27}, developmental switches controlled by ribosomal regulation might explain why mutations in ribosome biogenesis factors result in tissue-specific ribosomopathies^{2,19,20,28–30}. Manipulating such switches could lead to therapeutic strategies for paediatric diseases: as Treacher Collins syndrome is caused by mutation of a single TCOF1 allele, increasing the efficiency of KBTBD8-dependent ubiquitylation of the remaining wild-type TCOF1 might reconstitute ribosomal regulation and neural crest formation.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 26 March; accepted 24 July 2015.

- Gurdon, J. B. The egg and the nucleus: a battle for supremacy. *Development* **140**, 2449–2456 (2013).

2. The Treacher Collins Syndrome Collaborative Group. Positional cloning of a gene involved in the pathogenesis of Treacher Collins syndrome. *Nature Genet.* **12**, 130–136 (1996).
3. Dixon, J. *et al.* Tcof1/Treacle is required for neural crest cell formation and proliferation deficiencies that cause craniofacial abnormalities. *Proc. Natl Acad. Sci. USA* **103**, 13403–13408 (2006).
4. Petroski, M. D. & Deshaies, R. J. Function and regulation of cullin-RING ubiquitin ligases. *Nature Rev. Mol. Cell Biol.* **6**, 9–20 (2005).
5. Silverman, J. S., Skaar, J. R. & Pagano, M. SCF ubiquitin ligases in the maintenance of genome stability. *Trends Biochem. Sci.* **37**, 66–73 (2012).
6. Wang, Y. *et al.* Deletion of the *Cul1* gene in mice causes arrest in early embryogenesis and accumulation of cyclin E. *Curr. Biol.* **9**, 1191–1194 (1999).
7. Arai, T. *et al.* Targeted disruption of *p185/Cul7* gene results in abnormal vascular morphogenesis. *Proc. Natl Acad. Sci. USA* **100**, 9855–9860 (2003).
8. Singer, J. D., Gurian-West, M., Clurman, B. & Roberts, J. M. Cullin-3 targets cyclin E for ubiquitination and controls S phase in mammalian cells. *Genes Dev.* **13**, 2375–2387 (1999).
9. Jiang, B. *et al.* Lack of Cul4b, an E3 ubiquitin ligase component, leads to embryonic lethality and abnormal placental development. *PLoS ONE* **7**, e37070 (2012).
10. Voigt, J. & Papalopulu, N. A dominant-negative form of the E3 ubiquitin ligase Cullin-1 disrupts the correct allocation of cell fate in the neural crest lineage. *Development* **133**, 559–568 (2006).
11. Jin, L. *et al.* Ubiquitin-dependent regulation of COPII coat size and function. *Nature* **482**, 495–500 (2012).
12. Lander, R., Nordin, K. & LaBonne, C. The F-box protein Ppa is a common regulator of core EMT factors Twist, Snail, Slug, and Sip1. *J. Cell Biol.* **194**, 17–25 (2011).
13. Skaar, J. R., Pagan, J. K. & Pagano, M. Mechanisms and function of substrate recruitment by F-box proteins. *Nature Rev. Mol. Cell Biol.* **14**, 369–381 (2013).
14. Barbieri, C. E. *et al.* Exome sequencing identifies recurrent SPOP, FOXA1 and MED12 mutations in prostate cancer. *Nature Genet.* **44**, 685–689 (2012).
15. Louis-Dit-Picard, H. *et al.* KLHL3 mutations cause familial hyperkalemic hypertension by impairing ion transport in the distal nephron. *Nature Genet.* **44**, 456–460 (2012).
16. Welcker, M. & Clurman, B. E. FBW7 ubiquitin ligase: a tumour suppressor at the crossroads of cell division, growth and differentiation. *Nature Rev. Cancer* **8**, 83–93 (2008).
17. Yanai, I., Peshkin, L., Jorgensen, P. & Kirschner, M. W. Mapping gene expression in two *Xenopus* species: evolutionary constraints and developmental flexibility. *Dev. Cell* **20**, 483–496 (2011).
18. Chambers, S. M. *et al.* Highly efficient neural conversion of human ES and iPS cells by dual inhibition of SMAD signaling. *Nature Biotechnol.* **27**, 275–280 (2009).
19. Dauwerse, J. G. *et al.* Mutations in genes encoding subunits of RNA polymerases I and III cause Treacher Collins syndrome. *Nature Genet.* **43**, 20–22 (2011).
20. Jones, N. C. *et al.* Prevention of the neurocristopathy Treacher Collins syndrome through inhibition of p53 function. *Nature Med.* **14**, 125–133 (2008).
21. Price, J. C., Guan, S., Burlingame, A., Prusiner, S. B. & Ghaemmaghami, S. Analysis of proteome dynamics in the mouse brain. *Proc. Natl Acad. Sci. USA* **107**, 14508–14513 (2010).
22. Cambridge, S. B. *et al.* Systems-wide proteomic analysis in mammalian cells reveals conserved, functional protein turnover. *J. Proteome Res.* **10**, 5275–5284 (2011).
23. Tuck, A. C. & Tollervey, D. A transcriptome-wide atlas of RNP composition reveals diverse classes of mRNAs and lncRNAs. *Cell* **154**, 996–1009 (2013).
24. Jack, K. *et al.* rRNA pseudouridylation defects affect ribosomal ligand binding and translational fidelity from yeast to human cells. *Mol. Cell* **44**, 660–666 (2011).
25. Yoon, A. *et al.* Impaired control of IRES-mediated translation in X-linked dyskeratosis congenita. *Science* **312**, 902–906 (2006).
26. Kondrashov, N. *et al.* Ribosome-mediated specificity in Hox mRNA translation and vertebrate tissue patterning. *Cell* **145**, 383–397 (2011).
27. Xue, S. & Barna, M. Specialized ribosomes: a new frontier in gene regulation and organismal biology. *Nature Rev. Mol. Cell Biol.* **13**, 355–369 (2012).
28. McCann, K. L. & Baserga, S. J. Genetics. Mysterious ribosomopathies. *Science* **341**, 849–850 (2013).
29. Freed, E. F., Prieto, J. L., McCann, K. L., McStay, B. & Baserga, S. J. NOL11, implicated in the pathogenesis of North American Indian childhood cirrhosis, is required for pre-rRNA transcription and processing. *PLoS Genet.* **8**, e1002892 (2012).
30. Sondalle, S. B. & Baserga, S. J. Human diseases of the SSU processome. *Biochim. Biophys. Acta* **1842**, 758–764 (2014).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank R. Harland, J. Schatzky, R. Zoncu, J. Corn, A. Manford and E. Oh for advice and critical reading of the manuscript, and all members of the Rape and Ingolia laboratories for helpful discussions. A.W. was funded in part by a postdoctoral fellowship the California Institute of Regenerative Medicine; S.I. is supported by a Human Frontier Science Fellowship; C.A.M. was supported by an NSF graduate student fellowship; and S.M.-R. was supported by a UC Mexus-Conacyt fellowship and NIH GM42341. This work used the Vincent J. Coates Genomics Sequencing Laboratory and Proteomics/Mass Spectrometry Laboratory at UC Berkeley (NIH grants S10RR029668, S10RR027303 and S10RR025622). This work was funded by a Basic Biology Award from the California Institute of Regenerative Medicine to M.R. (RB-02222). M.R. is an investigator of the Howard Hughes Medical Institute.

Author Contributions A.W. designed and performed most experiments and helped to write the manuscript; S.I. performed ribosomal profiling and RNA processing experiments; C.A.M. and S.M.-R. performed *Xenopus* experiments; N.T. performed binding studies; I.F. wrote code for analysis of CompPASS mass spectrometry; N.T.I. supervised ribosomal profiling studies; M.R. helped design experiments and wrote the manuscript.

Author Information Raw data for microarray, RNA-seq and ribosome profiling was deposited in the NCBI Gene Expression Omnibus under accession numbers GSE62123 (microarray data) and GSE62247 (RNA seq and foot-printing data). Reprints and permissions information is available at www.nature.com/reprints. The authors declare competing financial interests: details are available in the online version of the paper. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to M.R. (mrpe@berkeley.edu).

METHODS

No statistical methods were used to predetermine sample size.

Plasmids, shRNAs, siRNAs, morpholinos. Full-length KBTBD8 was cloned into pCNA5 untagged, with a C-terminal $3 \times$ Flag tag, or a C-terminal HA tag, for expression in human cells, into pMAL with an N-terminal MBP tag for expression in *Escherichia coli*, and into pFastBac with an N-terminal $6 \times$ His tag for expression in Sf9 ES insect cells using the Bac-to-Bac baculovirus expression system (Invitrogen). For expression in human embryonic stem cells, KBTBD8 was cloned into pENTR1A with a C-terminal Flag tag and recombined into pLenti-PGK-Hygro. Point mutants in KBTBD8's BTB-BACK domain (Y74A, M78D, L73D, ST86,87AA, Y43A) and KBTBD8's propeller domain (I352A, F450A, MF496, 497AA, F550A, and W579A) were introduced into these vectors by site-directed mutagenesis and digestion of parental DNA with DpnI. Full-length coding sequences of *TCOF1* and *NOLC1* were cloned into pCMV- $3 \times$ Flag or pCS2-HA vectors for expression in human cells and *in vitro* transcription/translation (IVT/T), respectively, as described³¹. His-ubiquitin was cloned into a pCS2 vector for expression in cells³². To generate a deubiquitylation-resistant ubiquitin variant, a L73P mutation was introduced in this vector by site-directed mutagenesis as described above. pLKO1-Puro Mission shRNA constructs targeting *KBTBD8* (#1, TRCN0000130280; #2, TRCN0000128536), *NOLC1* (#1, TRCN000061971; #2, TRCN0000298197), *TCOF1* (#1, TRCN000008630), and *BRD2* (#1, TRCN0000066310; #2, TRCN0000381007), *DHX15* (#1, TRCN0000000006; #2, TRCN0000425479), *PCNT* (#1, TRCN0000162352; #2, TRCN0000162298) and *ANKUB1* (#1, TRCN0000336646) were purchased from Sigma. siRNAs were from Dharmacon and sequences were as follows: scrambled control (UAGCGACUAAACACA UCAAUU), *ARRB1* (#1, CGAGAGAAUUGGAGGAGAAUU; #2, UCAUAGA ACUUGACAAAUU), *ARRB2* (#1, ACAAGGAGGUGCUGGAAUUU; #2, CUAAAUCACUAGAAGAGAAUU), *NOLC1* (#1, CCAAGAAGGCUGUGGA GAAUU; #2, CAGUUAAGCUCAGAUAAUU; #3, UCUCAGAGGUGGCC AAUAAUU), *TCOF1* (#1, CCAUCAAGCAUGAAAGAAUU; #2, GGAAAGG CCUCCAGGUGAAUU; #3, GGAAUCAGACAGUGAGGAAUU).

Four morpholino oligonucleotides (Gene Tools, LLC) were used: *kbtbd8* translation blocking, 5'-GTGCAGGAAACGTCACCTACTTCTCT-3'; *kbtbd8* splice blocking, 5'-TCTCCAGCCCCAAACAACC-3'; *cul3* splice blocking, 5'-AAG TATCCTATGAGTCTCACCGGGA-3'; and a control tracer morpholino with 3'-fluorescein modification, 5'-CCTCTTACCTCAGTTACAATTTATA-3'. Dominant-negative CUL3 (N-terminal 250 amino acids of human CUL3) was cloned into pCS2+ and mRNA was synthesized for injection using an *in vitro* transcription system.

Proteins. KBTBD8, KBTBD8(Y74A) and KBTBD8(W579A) were purified from ES Sf9 cells 72 h after transduction. Lysates were prepared in 50 mM sodium phosphate, pH 8, 500 mM NaCl, and 10 mM imidazole by incubation with 200 μ g ml⁻¹ lysozyme and sonication. Lysates were cleared by centrifugation and incubated with Ni-NTA (QIAGEN) for 2 h at 4 °C. Beads were washed with lysis buffer containing 0.1% Triton. Proteins were eluted in 200 mM imidazole in 50 mM sodium phosphate, pH 8, 500 mM NaCl. TEV protease for His-tag removal was added and proteins were dialysed overnight into 50 mM Tris pH 8.0, 100 mM NaCl, 2 mM DTT. Proteins were further purified by molecular sieving over a preparative Superdex 200 column in 50 mM Tris pH 8.0, 100 mM NaCl, 2 mM DTT. KBTBD8 or KBTBD8 mutant fractions were concentrated, aliquoted, flash-frozen in liquid nitrogen, and stored at -80 °C.

MBP-tagged KBTBD8, KBTBD8(Y74A), KBTBD8(F550A) and KBTBD8(W579A) were expressed and purified from BL21/DE3 (RIL) cells. Cells were grown in LB-medium to OD_{600 nm} 0.5 followed by addition of 0.5 mM IPTG and protein production for 16 h at 18 °C. Cells were harvested, lysed in 50 mM sodium phosphate, pH 8, 500 mM NaCl, by incubation with 200 μ g ml⁻¹ lysozyme and sonication, and cleared by centrifugation. MBP-tagged proteins were isolated using amylose resin (NEB), washed with lysis buffer, and eluted in lysis buffer containing 10 mM maltose followed by dialysis into PBS containing 2 mM DTT.

Antibodies. Mouse monoclonal anti-KBTBD8 antibodies were produced with Promab using MBP-KBTBD8-250-601 as antigen (1 μ g ml⁻¹ in immunoblots (IB) and 10 μ g mg⁻¹ lysate in immunoprecipitations (IP)). Anti-PAX6 (DSHB, clone P3U1, 1:1,000 in IB), anti-PAX6 (Biolegend, PRB-278B, 1:300 in IF), anti-TDGF (#4193, clone D81B12, Cell Signaling, 1:1,000 in IB), anti-TFAP2 (#2509, Cell Signaling, 1:1,000 in IB, 1:100 in IF), anti-CDH1 (#3195, clone 24E10, Cell Signaling, 1:1,000 in IB), anti-FOXG1 (ab18259, Abcam, 1:250 in IB), anti-p75 (AB-N07, clone ME20.4, Advanced Targeting Systems, 1:100 in IF), anti-HNK1 (C6680, clone VC1.1, Sigma, 1:500 in IF), anti-SOX10 (ab155279, Abcam, 1:2,000 in IB), anti-SOX10 (ab27655, Abcam, 1:100 in IF), anti-GFAP (#3670, clone GA5, Cell Signaling, 1:100 in IF), anti-SMA (A2547, clone 1A4, 1:100 in IF), anti-NF-L (#2837, clone C28E10, Cell Signaling, 1:100), anti-Tuj1 (#5568, clone D71G9, Cell

Signaling, 1:100 in IF), anti-rRNA 5.8 (ab347144, clone Y10b, Abcam, 1:1,000 in IF), anti-TCOF1 (11003-1-AP, Proteintech, 1:250 in IB), anti-TCOF1 (sc-49529, Santa Cruz, 1:100 in IF), anti-FBL (ab5821, Abcam, 1:100 in IF), anti-NOLC1 (11815-1-P, Proteintech, 1:1,000 in IB), anti-ARRB1/2 (#4674, clone D24H9, Cell Signaling, 1:1,000 in IB), anti-PKN1 (610687, clone 49/PRK1, BD Bioscience, 1:1,000 in IB), anti-CUL3 (Bethyl, 1:1,000 in IB), anti-NANOG (#3580, Cell Signaling, 1:1,000 in IB), anti-OCT4 (ac-8628, Santa Cruz, 1:1,000 in IB), anti-SNAIL2 (#9585, clone C19G7, Cell Signaling, 1:500 in IB), anti-GAPDH (#2118, clone 14C10, Cell Signaling, 1:10,000 in IB), anti-BRD2 (#5848, clone D89B4, Cell Signaling, 1:500 in IB), anti-PCNA (Santa Cruz, 1:5,000 in IB), anti-RPA1 (Santa Cruz, 1:100 in IB, 4 μ g per 1 mg lysate in IP), anti-RPA2 (sc-17913, Santa Cruz, 1:100 in IB), anti-RPB1 (sc-5943, Santa Cruz, 1:100 in IB), anti-DKC1 (sc-48794, Santa Cruz, 1:1,000 in IB), anti-NHP2 (sc-366967, Santa Cruz, 1:500 in IB), anti-NOP56 (A302-720A, Bethyl, 1:1,000 in IB), anti-NOP58 (A302-719A, Bethyl, 1:1,000 in IB), anti-CSK2A (#2656, Cell Signaling, 1:500 in IB), anti-cleaved caspase-3 (#9664, clone 5A1E, Cell Signaling, 1:100 in IB), anti-ATRX (A301-045A, Bethyl, 1:500 in IB), anti-PCM1 (A301-149A, Bethyl, 1:500 in IB), anti-SNF2L (#12438, clone D4Q7V, Cell Signaling, 1:500 in IB), anti-HA (clone C29F4; Cell Signaling, 1:3,000 in IB), anti-Flag (F1804, clone M2, Sigma, 1:2,000 in IB), and anti-Flag (F7425, Sigma, 1:2,000 in IB) antibodies were commercially purchased. **Mammalian cell culture and transfections.** Human embryonic kidney (HEK) 293T cells were maintained in DMEM with 10% fetal bovine serum. Plasmid transfections of HEK 293T cells were with calcium phosphate and siRNA transfections were with Lipofectamine RNAiMAX (Invitrogen) according to the manufacturer's instructions using 10 nM for each siRNA.

hES cell culture, lentiviral infections and hES differentiations. Human embryonic stem (hES) H1 cells were obtained from the Wisconsin stem cell bank, routinely characterized for mycoplasma contamination, and maintained under feeder free conditions on Matrigel-coated plates (#354277, BD Biosciences) in mTeSR1, (#05871/05852, StemCell Technologies Inc.) and were routinely passaged with collagenase (#07909, StemCell Technologies Inc.) and ReLeSR (#05872, StemCell Technologies Inc.).

Lentiviruses were produced in 293T cells by co-transfection of lentiviral constructs with packaging plasmids (Addgene) for 48–72 h. Transduction was carried out by infecting 30% confluent hES H1 cells with lentiviruses in the presence of 6 μ g ml⁻¹ Polybrene (Sigma). After 7 days of selection with appropriate antibiotic (0.5 μ g ml⁻¹ puromycin for pLKO1-puro-shRNA constructs, 500 μ g ml⁻¹ hygromycin for pLKO1-hygro-shRNA or pLenti-Hygro constructs), hES H1 cells were analysed and used in differentiation experiments.

Embryoid body (EB) formation from hES H1 cells and hES H1 cells expressing control shRNA or shRNAs targeting KBTBD8 was performed using Aggrewell 800 plates (#27865, StemCell Technologies Inc.) and APEL medium (#05210, StemCell Technologies Inc.) following the guidelines of the manufacturer's technical manual (#29146). In brief, single-cell suspensions were prepared by treatment of hES cells with accutase (#07920, StemCell Technologies Inc.) and 1×10^6 cells were seeded per well of an Aggrewell 800 plate in APEL medium supplemented with 10 μ M Y-27632 ROCK inhibitor (Calbiochem). 24 h after seeding, EBs were harvested and transferred into ultra-low adherence culture dishes (Corning) with < 1,000 EBs per well of a 6-well plate and differentiated in APEL medium for 3 and 6 days. Medium was replaced every other day.

Neural induction of hES H1 cells expressing different shRNA constructs was performed using STEMdiff Neural Induction Medium (#05831, StemCell Technologies Inc.) in combination with a monolayer culture method according to the manufacturer's technical bulletin (#28044) and as previously described¹⁸. In brief, single-cell suspensions were prepared by treatment of hES cells with accutase and $1.25\text{--}1.5 \times 10^6$ cells were seeded per well of a 6-well plate in STEMdiff Neural Induction Medium supplemented with 10 μ M Y-27632 ROCK inhibitor. Neural induction was performed for 1, 3, 6 and 9 days with daily medium change. For inhibition of general protein translation, hESCs were subjected to the neural conversion protocol in the presence of 50 or 100 nM rapamycin.

For long-term neural conversion experiments to assess spontaneous differentiation of neural crest cells into derivatives, hESCs were subjected to neural conversion using STEMdiff Neural Induction Medium as described above for 43 days. Medium was changed daily until 18 days, then every other day.

Microinjections and *in situ* hybridizations. Morpholinos (20 ng) and mRNA (100 pg) with 2 ng of tracer morpholino were injected into the animal cap in 1 blastomere of 2-cell stage *Xenopus tropicalis* embryos. At stage 10–14, embryos were sorted by left or right injection side via tracer morpholino fluorescence. Embryos were developed to stage 16–18 and fixed for 4–6 h in MEMFA at room temperature. *In situ* hybridization of *X. tropicalis* embryos with digoxigenin-labelled RNA probes was performed using a multi-basket method as described previously³³. Sorting and imaging were performed on a Zeiss SteREO Lumar.V12 microscope.

Cycloheximide (CHX) chase assays. For cycloheximide chase assays, control or KBTBD8-depleted hES H1 cells and cells that had undergone neural induction for 3 days were treated with 40 $\mu\text{g ml}^{-1}$ CHX for 2, 4 and 6 h. Cells were lysed in 8 M urea, 50 mM sodium phosphate, pH 8.0. Lysates were diluted in SDS loading buffer, sonicated, and were analysed by immunoblotting. For quantification, immunoblot signals for respective proteins were quantified using ImageJ (NIH, <http://rsbweb.nih.gov/ij/>) and normalized to GAPDH or β -actin.

Gene expression analysis by microarray. To compare gene expression profiles of hES H1 cells versus embryoid bodies (6 days), gene expression profiles of control versus KBTBD8-depleted hES H1 cells, and gene expression profiles of control versus KBTBD8-depleted embryoid bodies (6 days), we isolated and purified total RNA from respective samples using the RNeasy Mini Kit (Qiagen, catalogue no. 74104). Microarray analysis was performed in biological triplicates by the Functional Genomics Laboratory (UC Berkeley) using the Affymetrix HUMAN GENE 1.0 ST ARRAY.

Quantitative real-time PCR (qRT-PCR) analysis. For qRT-PCR analysis, total RNA was extracted and purified from cells using the RNeasy Mini Kit (Qiagen, catalogue no. 74104) and transcribed into cDNA using the RevertAid first strand cDNA synthesis kit (#K1621, Thermo Scientific). Gene expression was quantified by Maxima SYBR Green/Rox qPCR (#K0221, Thermo scientific) on a StepOnePlus Real-Time PCR System (Applied Biosystems). Nonspecific signals caused by primer dimers were excluded by dissociation curve analysis and use of non-template controls. To normalize for loaded cDNA, β -actin or RPS6 was used as endogenous control. Gene-specific primers for qRT-PCR were designed by using NCBI Primer-Blast or ordered pre-designed from Integrated DNA Technologies. Primer sequences can be found in Supplementary Table 2.

Cluster analysis. To determine shRNA treatments that caused similar effects on neural conversion of hESCs, we performed cluster analysis of mRNA expression profiles. We stably transduced H1 hESCs with lentiviruses expressing various shRNAs and/or shRNA-resistant cDNAs and subjected these cells to neural conversion by dual SMAD inhibition for 9 days. We then measured the mRNA abundance of neural progenitor and neural crest markers by RT-qPCR. Data sets were clustered using the heatmap.2 function of the gplots package on R.

In vivo ubiquitylation assays. To detect ubiquitylation of ectopic TCOF1 or NOLC1, HEK 293T cells were transiently transfected with 6 \times His-tagged ubiquitin and HA-tagged TCOF1 or NOLC1. For detection of ubiquitylation of endogenous TCOF1 or NOLC1, HEK 293T cells were transiently transfected with 6 \times His-tagged ubiquitin-L73P. Cells were harvested, washed with PBS, lysed in 8 M urea, 50 mM sodium phosphate, pH 8.0 and sonicated. His-ubiquitin conjugates were purified using Ni-NTA agarose (Qiagen) and ubiquitylated NOLC1 or TCOF1 was detected by immunoblotting using anti-HA, anti-TCOF1, or anti-NOLC1 antibodies. For analysis of the influence of β -arrestin proteins on KBTBD8-mediated ubiquitylation, HEK 293T cells were transfected with control siRNAs or a pool of siRNAs targeting ARRB1 (#1 and #2) and ARRB2 (#1 and #2) 24 h before plasmid transfection.

Immunoprecipitations for mass spectrometry. Anti-Flag immunoprecipitations (IPs) for mass spectrometry analysis were performed from extracts of HEK 293T cells transiently expressing KBTBD8-3 \times Flag versions or 3 \times Flag-TCOF1 in the presence and absence of KBTBD8 versions (20 \times 15 cm dishes per condition). Lysis was in two pellet volumes of 20 mM HEPES pH 7.3, 150 mM NaCl, 110 mM KOAc, 2 mM Mg(OAc)₂, 5 mM EDTA, 5 mM EGTA, 0.2% NP-40, and protease inhibitors (Roche) on ice. Lysates were sonicated, cleared by centrifugation, passed through a 0.45 μm membrane filter, and incubated with anti-Flag-M2 agarose (Sigma) for 2 h at 4 $^{\circ}\text{C}$. After washing with lysis buffer, Flag-tagged protein complexes were eluted with lysis buffer containing 0.5 mg ml^{-1} 3 \times Flag peptide in three 15 min incubations at 30 $^{\circ}\text{C}$, 800 rpm. Eluates were either analysed by immunoblotting or further processed for multi-dimensional protein identification technology (MUDPIT) mass spectrometry. Sequential NOLC1-TCOF1 IPs were performed from extracts of HEK 293T cells transiently expressing 3 \times HA-NOLC1, 3 \times Flag-TCOF1, and untagged KBTBD8 (120 \times 15 cm plates for mass spectrometry). Lysates were prepared as described above and incubated with anti-HA agarose (Sigma) for 2 h at 4 $^{\circ}\text{C}$. After washing with lysis buffer, bound proteins were eluted with lysis buffer containing 0.5 mg ml^{-1} 3 \times HA peptide in three 15 min incubations at 30 $^{\circ}\text{C}$, 800 rpm. HA eluates were subjected to anti-Flag immunoprecipitation followed by Flag peptide elution (see above).

Endogenous immunoprecipitations. Anti-KBTBD8 immunoprecipitations were performed from hES H1 cells (5 \times 15 cm dishes per condition) and lysates were prepared as described above. After incubation with KBTBD8 or control antibodies (mIgGs, Santa Cruz) at 4 $^{\circ}\text{C}$ for 1 h, protein G beads (Roche) were added for 2 h. After washing with lysis buffer, bound proteins were eluted with 2 \times SDS sample buffer and analysed by immunoblotting. anti-RPA1 IPs were performed in the same way but from either HEK 293T cells transiently expressing

KBTBD8-Flag, KBTBD8(W579A)-Flag, or KBTBD8(Y74A)-Flag (2 \times 15 cm plates for each condition) or from control or KBTBD8-depleted hES H1 cells (5 \times 15 cm plates for each condition).

Mass spectrometry and compPASS analysis. For mass spectrometry analysis, Flag immunoprecipitates were prepared as described above and precipitated with 20% trichloroacetic acid (TCA, Fisher) overnight. Proteins were resolubilized and denatured in 8 M urea (Fisher), 100 mM Tris (pH 8.5), followed by reduction with 5 mM TCEP (Sigma), alkylation with 10 mM iodoacetamide (Sigma), and overnight digestion with trypsin (0.5 mg ml^{-1} , Fisher). Samples were analysed by the Vincent J. Coates Proteomics/Mass Spectrometry Laboratory at UC Berkeley and compared with \sim 100 reference immunoprecipitations against different Flag-tagged bait proteins using a Java script programmed according to the CompPASS software suite³⁴. For determination of the KBTBD8 interaction network, three independent KBTBD8-Flag IPs were compared as replicates against the reference IPs. Thresholds for high confidence interaction partners (HCIPs) were top 5% of interactors with highest Z-score and highest WD score. To narrow down putative substrates of KBTBD8 in the interaction map, we compared relative total spectral counts for each HCIP found in wild-type KBTBD8 immunoprecipitates to the ones found in KBTBD8(Y74A), KBTBD8(F550A) and KBTBD8(W579A) immunoprecipitates. For identification of effector proteins recruited to TCOF1 upon ubiquitylation, we determined the TCOF1 interaction network as described above for KBTBD8 and compared relative total spectral counts for each TCOF1 HCIP to those found upon co-expression of KBTBD8 or KBTBD8(Y74A). We then plotted relative TSC changes upon KBTBD8 expression against the difference of relative TSC changes upon KBTBD8 and KBTBD8(Y74A) expression.

Immunofluorescence microscopy. For immunofluorescence analysis, hES H1 cells or hES H1 cells expressing different shRNA were seeded on Matrigel-coated coverslips using accutase, fixed with 3.7% formaldehyde for 10 min, permeabilized with 0.1% Triton for 20 min, and stained with indicated antibodies or Hoechst. Images were taken using Zeiss LSM 710 confocal microscope or Olympus IX81 microscope, deconvolved using Metamorph, and processed using ImageJ.

Determination of average nucleolar size. To analyse nucleolar integrity, we performed indirect immunofluorescence microscopy using antibodies against fibrillarin, an established nucleolar marker. We stained nucleoli of control and KBTBD8-depleted hESCs or of hESCs that were subjected to neural conversion for 3 days. Images were taken for each condition using a Zeiss LSM 710 confocal microscope with a 20 \times objective followed by quantification of average nucleolar and nuclear size using ImageJ. Average nucleolar size was expressed relative to average nuclear area. Error bars represent standard deviation of three different images (\sim 100 cells per image).

Analysis of cell cycle progression. For DNA content analysis, control or KBTBD8-depleted hES H1 cells or control or KBTBD8-depleted cells that had undergone neural induction for 3 or 6 days were fixed in 70% ethanol in PBS overnight. Cells were pelleted and resuspended in PBS containing 1 mg ml^{-1} RNase (Sigma) and 10 $\mu\text{g ml}^{-1}$ propidium iodide (PI), incubated at room temperature for 30 min, then analysed using a Beckman-Coulter EPICS XL Flow Cytometer (575 nm band pass filter).

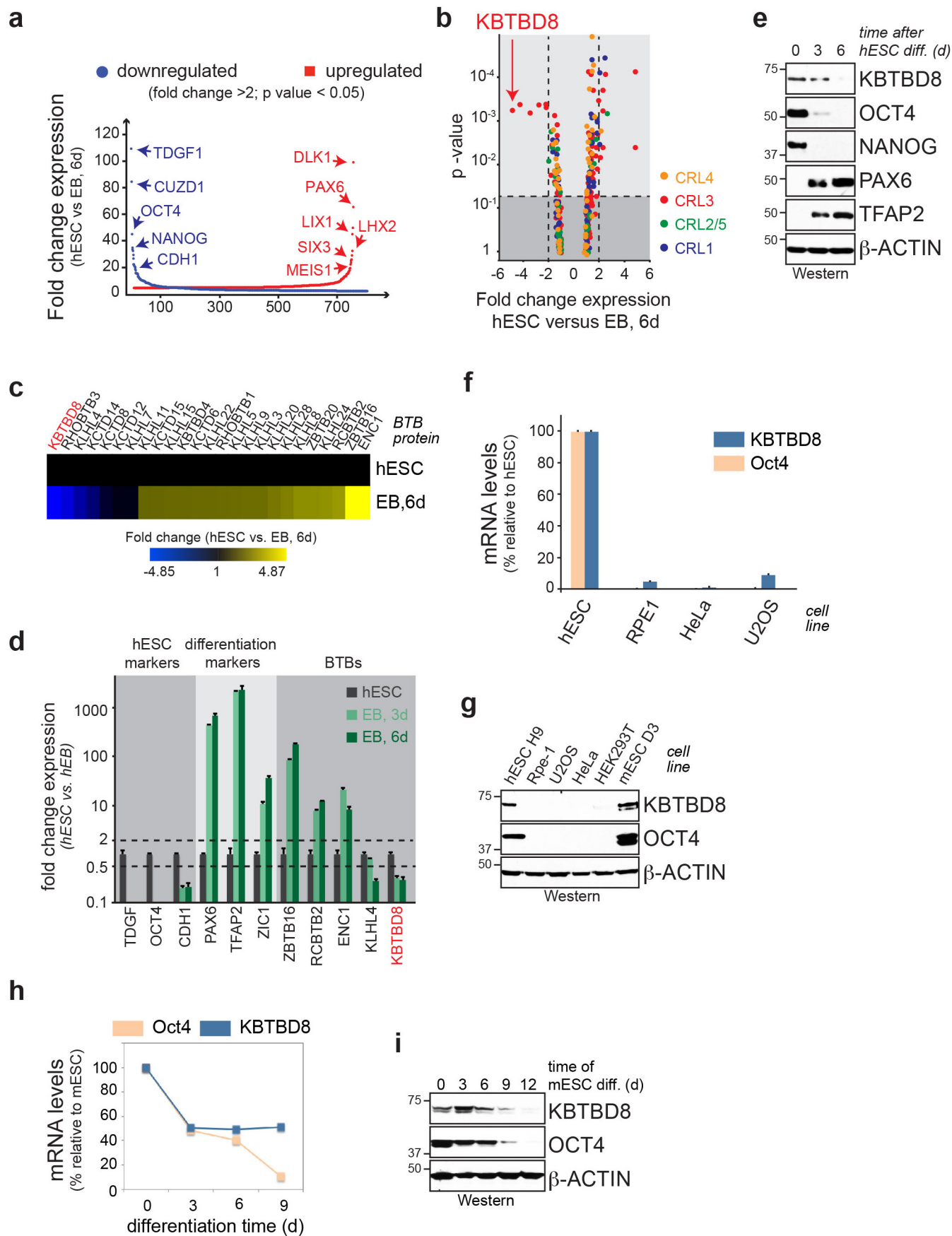
Cell proliferation assays. To compare the division rate of control and KBTBD8-depleted hES cells, we seeded 3 \times 10⁵ cells per well of a 6-well plate using accutase. Cells were accutased at 2, 3 and 4 days post seeding and counted using a haemocytometer.

RNA sequencing and ribosome profiling. RNA-seq libraries were prepared with Tru-seq Ribo-zero gold kit (Illumina). The preparation of ribosome profiling library and the data analysis were performed according to the method previously described³⁵. The libraries were sequenced on a HiSeq 2000 (Illumina). The reads were aligned to the hg19 human genome reference and the resulting aligned reads were mapped to UCSC known reference genes. Based on length of each footprint, specific A-site offsets were estimated as 14 for 26–28 nucleotides and 15 for 29–31 nucleotides. For mRNA fragments, we used offset 14. For measuring footprint density and mRNA fragments between samples, we restricted our analysis to genes that have at least 128 summed counts in each sample, only including the genomic positions 15 codons following the start codon and the position 5 codons preceding the stop codon. DESeq³⁶ was used to calculate fold change enrichment of genes by KBTBD8 or TCOF1 knockdown at each time point after neural induction.

Polysome profiling. Cells lysate was prepared as described³⁵. Lysate containing 3 μg total RNA was loaded on to 10–50% linear sucrose gradients containing 20 mM Tris pH 7.4, 150 mM NaCl, 5 mM MgCl₂, 1 mM DTT, 100 $\mu\text{g ml}^{-1}$ cycloheximide, and 2 U ml⁻¹ SUPERase in RNase inhibitor and centrifuged at 36,000 rpm for 2.5 h at 4 $^{\circ}\text{C}$ with SW41 rotor (Beckman Coulter). UV absorbance of fractionated gradient with Gradient station (Biocomp) was detected by ECONO UV monitor (Biorad). Monosome/Polysome ratio was determined by integration of the area under the respective peaks using Igor Pro software (WaveMetrics).

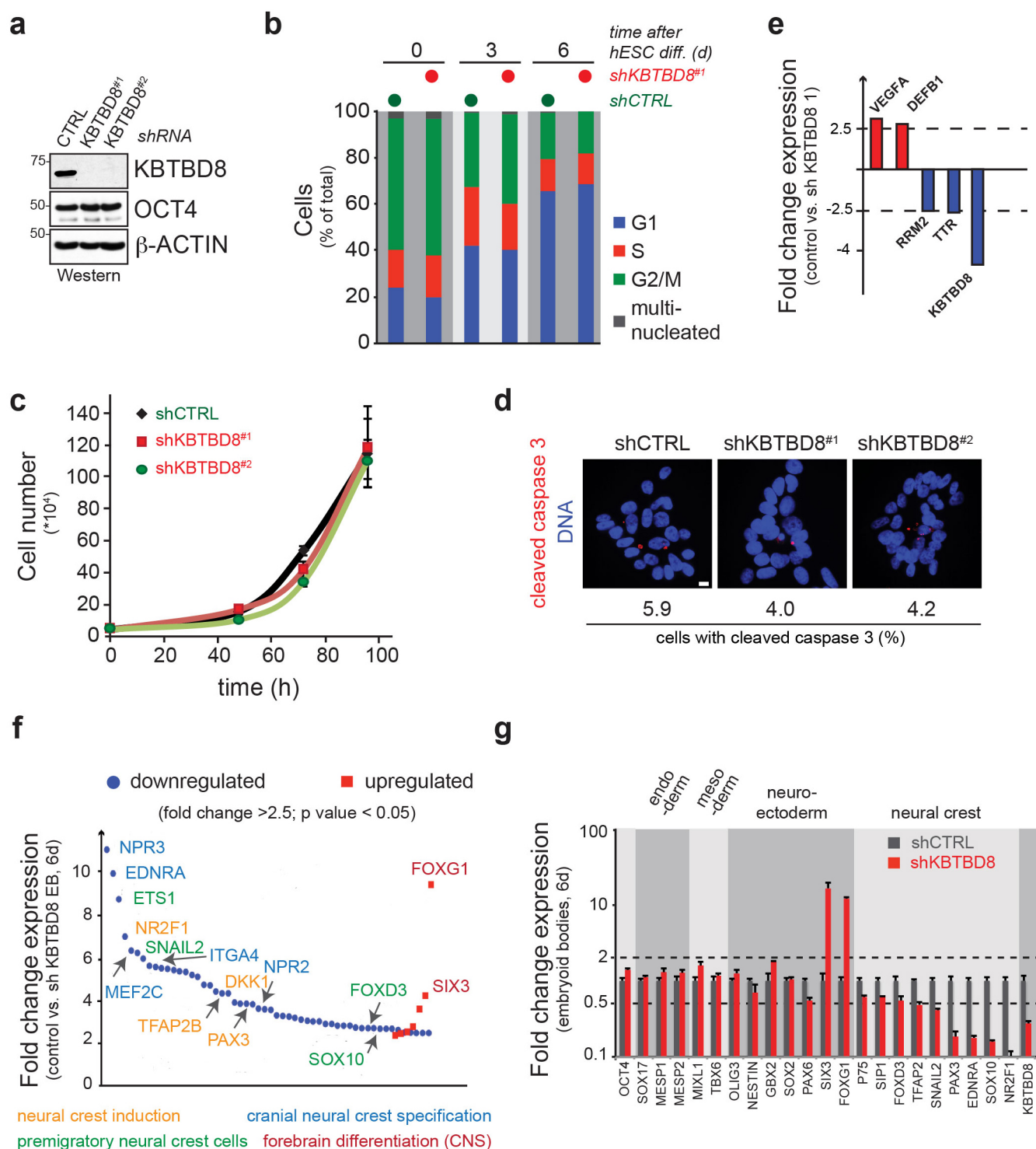
Metabolic labelling. To determine global mRNA translation in control or KBTBD8-depleted hESCs, hESCs subjected to embryoid body formation for 3 days, or hESCs subjected to neural conversion for 3 days, we employed metabolic labelling. Cells were pre-equilibrated in labelling medium (RPMI without methionine containing 10% dialysed FBS) for 15 min, followed by preparation of single-cell suspensions by accutase treatment. 0.6×10^6 cells for hESCs and hESCs that were subjected to neural conversion for three days, and ~ 300 EBs (initial seeding cell number: 3,000 cells per EB) were pulsed with 100 μ l labelling medium containing 0.1 mCi ml⁻¹ [³⁵S]-methionine at 37 °C for 20 min. After washing with PBS, cells were lysed in 8 M urea, 100 mM Tris pH 8.0, diluted in SDS loading buffer, sonicated, and lysates were analysed by SDS-PAGE and autoradiography. [³⁵S]-methionine incorporation was quantified using Image J software, normalized to total protein amount, and expressed relative to the control hESC sample (set to 1). Error bars denote standard deviation of three biological replicates.

31. Rape, M., Reddy, S. K. & Kirschner, M. W. The processivity of multiubiquitination by the APC determines the order of substrate degradation. *Cell* **124**, 89–103 (2006).
32. Jin, L., Williamson, A., Banerjee, S., Philipp, I. & Rape, M. Mechanism of ubiquitin-chain formation by the human anaphase-promoting complex. *Cell* **133**, 653–665 (2008).
33. Khokha, M. K. *et al.* Techniques and probes for the study of *Xenopus tropicalis* development. *Develop. Dyn.* **225**, 499–510 (2002).
34. Sowa, M. E., Bennett, E. J., Gygi, S. P. & Harper, J. W. Defining the human deubiquitinating enzyme interaction landscape. *Cell* **138**, 389–403 (2009).
35. Ingolia, N. T., Brar, G. A., Rouskin, S., McGeachy, A. M. & Weissman, J. S. The ribosome profiling strategy for monitoring translation *in vivo* by deep sequencing of ribosome-protected mRNA fragments. *Nature Protocols* **7**, 1534–1550 (2012).
36. Anders, S. & Huber, W. Differential expression analysis for sequence count data. *Genome Biol.* **11**, R106 (2010).



Extended Data Figure 1 | KBTBD8 is a developmentally regulated CUL3 adaptor. **a**, Gene expression analysis by microarray of hESCs differentiated into embryoid bodies (EB) for 6 days ($n > 30,000$ transcripts, mean of 3 biological replicates, analysis of variance (ANOVA) P value < 0.05 ; blue, downregulated genes; red, upregulated genes). **b**, Expression analysis of all CRL substrate adaptors, including KBTBD8, with data derived from the experiment described above. **c**, Expression analysis of CUL3 adaptors during hESC differentiation into hEBs (blue, downregulation; yellow, upregulation). **d**, mRNA levels of pluripotency markers and KBTBD8 during hESC differentiation into EBs, as determined by qRT-PCR (mean of 3 technical

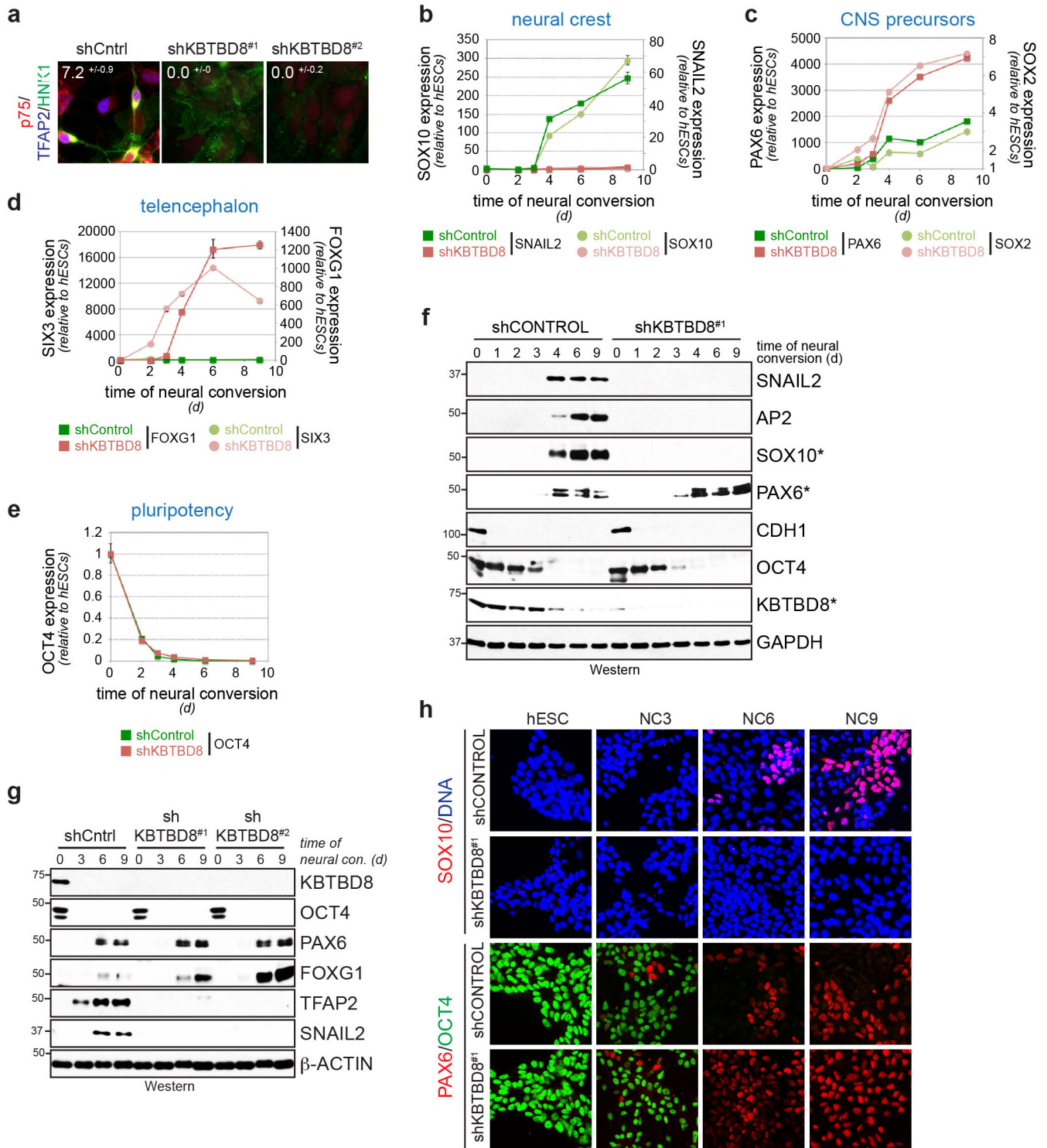
replicates \pm s.e.m.). **e**, Protein levels of KBTBD8 during hESC differentiation into EBs, as seen by western blot (OCT4, NANOG: pluripotency; PAX6: CNS precursors; TFAP2: neural crest marker). **f**, KBTBD8 is expressed in hESCs, but not in somatic cell lines, as determined by qRT-PCR (mean of 3 technical replicates \pm s.e.m.). **g**, Abundance of KBTBD8 in H9 hESCs, D3 mESCs, or somatic cell lines was determined by western blot analysis. **h**, KBTBD8 expression is downregulated during mouse embryonic stem cell (mESC) differentiation into mouse embryoid bodies, as determined by qRT-PCR (mean of 3 technical replicates \pm s.e.m.). **i**, KBTBD8 protein levels are reduced during mESC differentiation, as shown by western blot analysis.



Extended Data Figure 2 | KBTBD8 controls neural crest formation.

a, Stable depletion of KBTBD8 from H1 hESCs, as determined by western blot analysis. **b**, KBTBD8 depletion does not significantly change the cell cycle profile of hESCs, as determined by propidium iodide staining and FACS. **c**, Control or KBTBD8-depleted hESCs were counted at indicated times after seeding (mean of 3 biological replicates, \pm s.d.). **d**, KBTBD8 depletion does not induce apoptosis in hESCs, as shown by immunostaining against cleaved caspase 3 (red) or DNA (blue) (200 cells per condition; scale bar, 10 μ m). **e**, KBTBD8 depletion does not affect the gene expression profile of hESCs, as

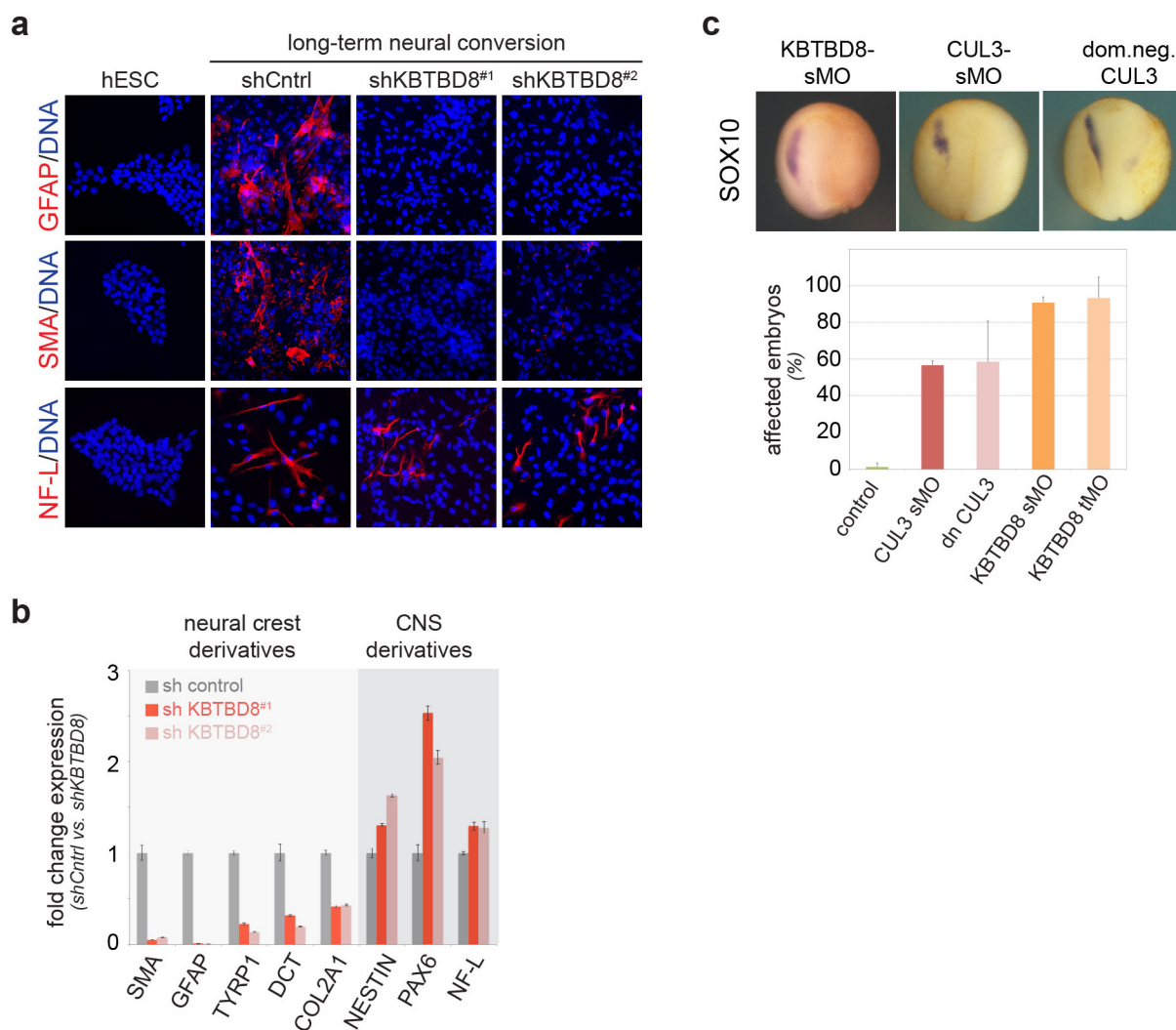
determined by microarray analysis (genes > 2.5 -fold change, $n > 30,000$; mean of 3 biological replicates, ANOVA P -value < 0.05). **f**, Loss of KBTBD8 causes a decrease in the expression of neural crest cell markers during EB formation, as shown by comparative microarray analysis (genes > 2.5 -fold change, $n > 30,000$; mean of 3 biological replicates, ANOVA P -value < 0.05). **g**, mRNA levels of pluripotency and differentiation markers in EBs stably expressing control or KBTBD8 shRNAs were measured by qRT-PCR (3 technical replicates \pm s.e.m.).



Extended Data Figure 3 | KBTBD8 controls neural crest specification.

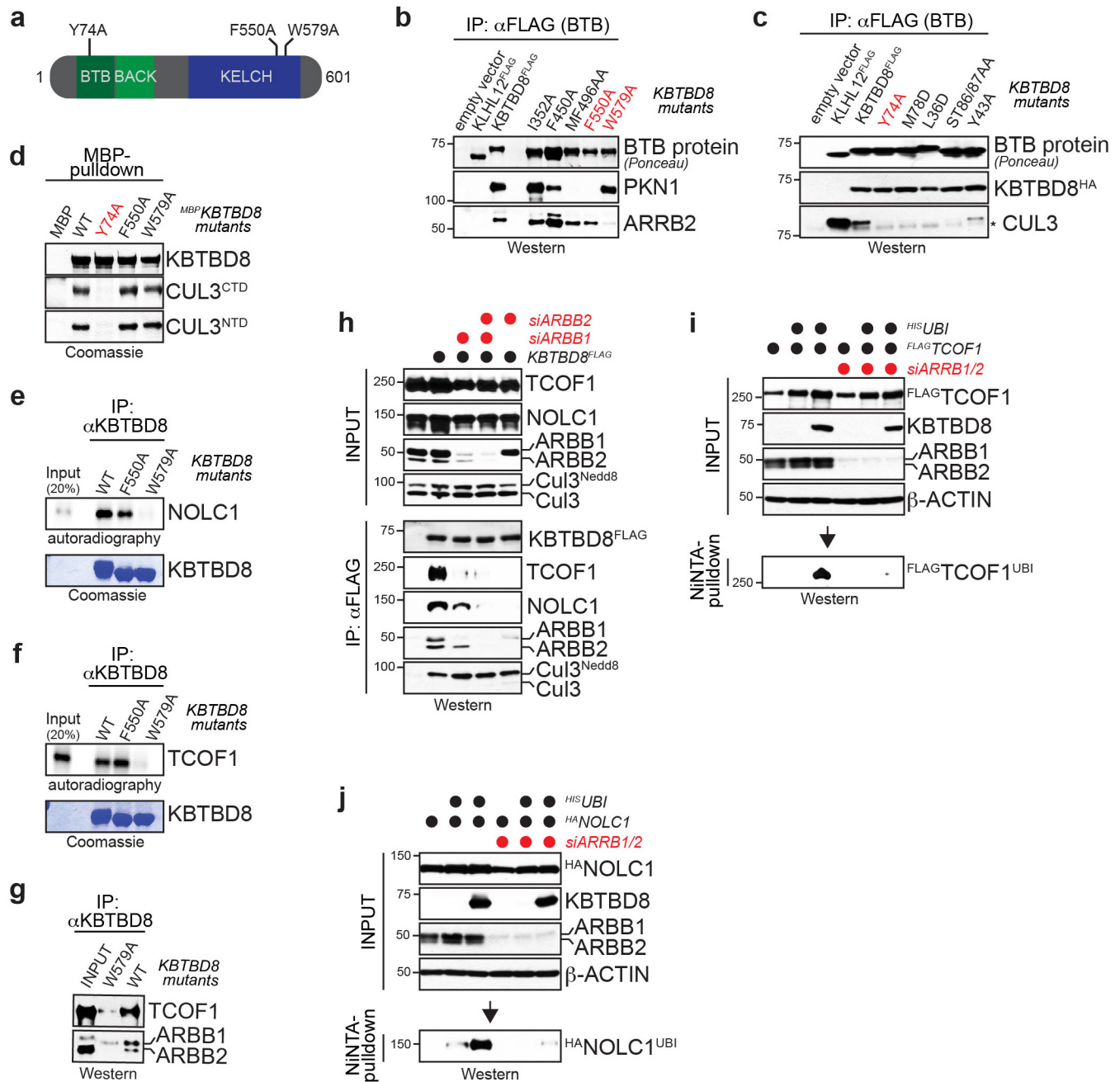
a, Depletion of KBTBD8 from hESCs subjected to neural conversion results in loss of neural crest cells, as determined by immunofluorescence against HNK1, TFAP2 and p75 ($n > 200$ cells, mean of 3 biological replicates \pm s.d.). **b**, H1 hESCs transduced with control (green) or *KBTBD8* shRNAs (red) were subjected to neural conversion, and expression of neural crest markers SOX10 (circles) and SNAIL2 (boxes) was monitored by qRT-PCR (mean of 3 technical replicates \pm s.e.m.). **c**, H1 hESCs described above were subjected to neural conversion, and abundance of CNS precursor markers SOX2 (circles) and PAX6 (boxes) was measured by qRT-PCR. **d**, H1 hESCs described above were subjected to neural conversion, and abundance of telencephalon markers SIX3 (circles) and FOXG1 (boxes) was measured by qRT-PCR. **e**, Expression of

OCT4 was monitored by qRT-PCR during neural conversion in the presence or absence of KBTBD8. **f**, hESCs stably expressing control or *KBTBD8* shRNAs were subjected to neural conversion and analysed for expression of pluripotency (OCT4, CDH1), neural crest (SOX10, SNAIL2, AP2), or CNS precursor markers (PAX6) by western blotting. To provide consistency, samples were taken from the same experiment as shown in Fig. 5d (asterisks mark blots that are also shown in Fig. 5d). **g**, Loss of neural crest occurs in response to KBTBD8 depletion by two independent shRNAs, as shown by western blot analysis. **h**, hESCs were subjected to neural conversion and analysed by immunofluorescence microscopy against SOX10 (neural crest), PAX6 (CNS precursor), and OCT4 (pluripotency) (confocal, original magnification 20 \times).



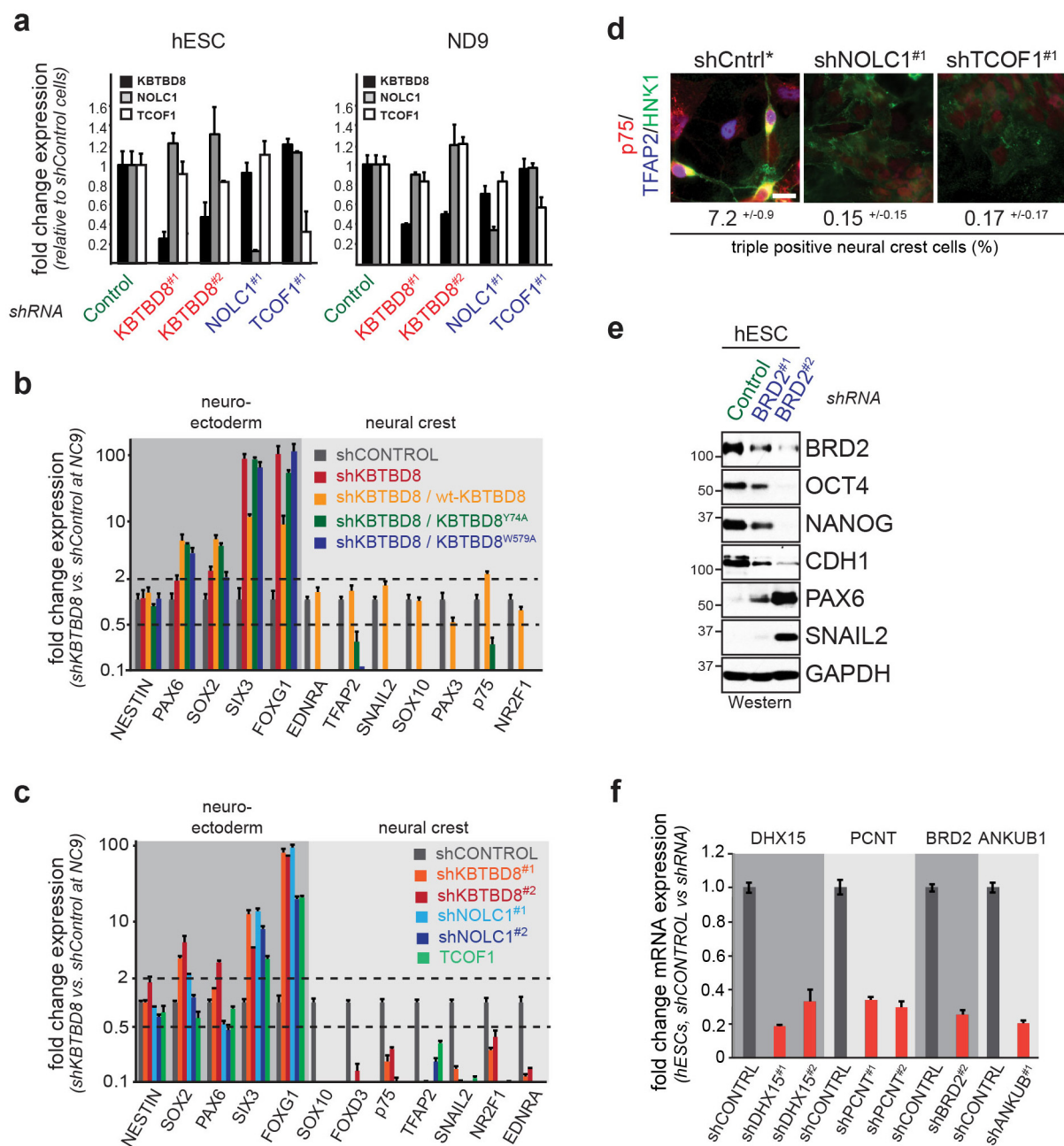
Extended Data Figure 4 | KBTBD8 is required for differentiation into functional neural crest cells. **a**, H1 hESCs stably expressing control or *KBTBD8* shRNAs were subjected to neural conversion for 43 days and analysed by immunofluorescence microscopy against GFAP (glia), smooth muscle actin (SMA; mesenchymal cells), and neurofilament L (neurons). **b**, Control H1 hESCs or hESCs depleted of *KBTBD8* were subjected to neural conversion for 43 days and expression of markers for glia (GFAP), mesenchyme (smooth muscle actin, SMA), melanocytes (TYRP1, DCT), chondrocytes (COL2A1), or

CNS derivatives (PAX6, NESTIN, neurofilament L) was analysed by qRT-PCR (mean of 3 technical replicates \pm s.e.m.). **c**, *Xenopus tropicalis* embryos were injected at the two-cell stage with splice-blocking morpholinos (sMO) against CUL3 or KBTBD8, or with a dominant-negative construct of CUL3 that allows KBTBD8 to bind, but not ubiquitylate, substrates. Neural crest formation was monitored by SOX10 *in situ* hybridization. Quantification included experiment shown in Fig. 1d (mean of 3 biological replicates \pm s.d.; ~20 embryos per condition and replicate).



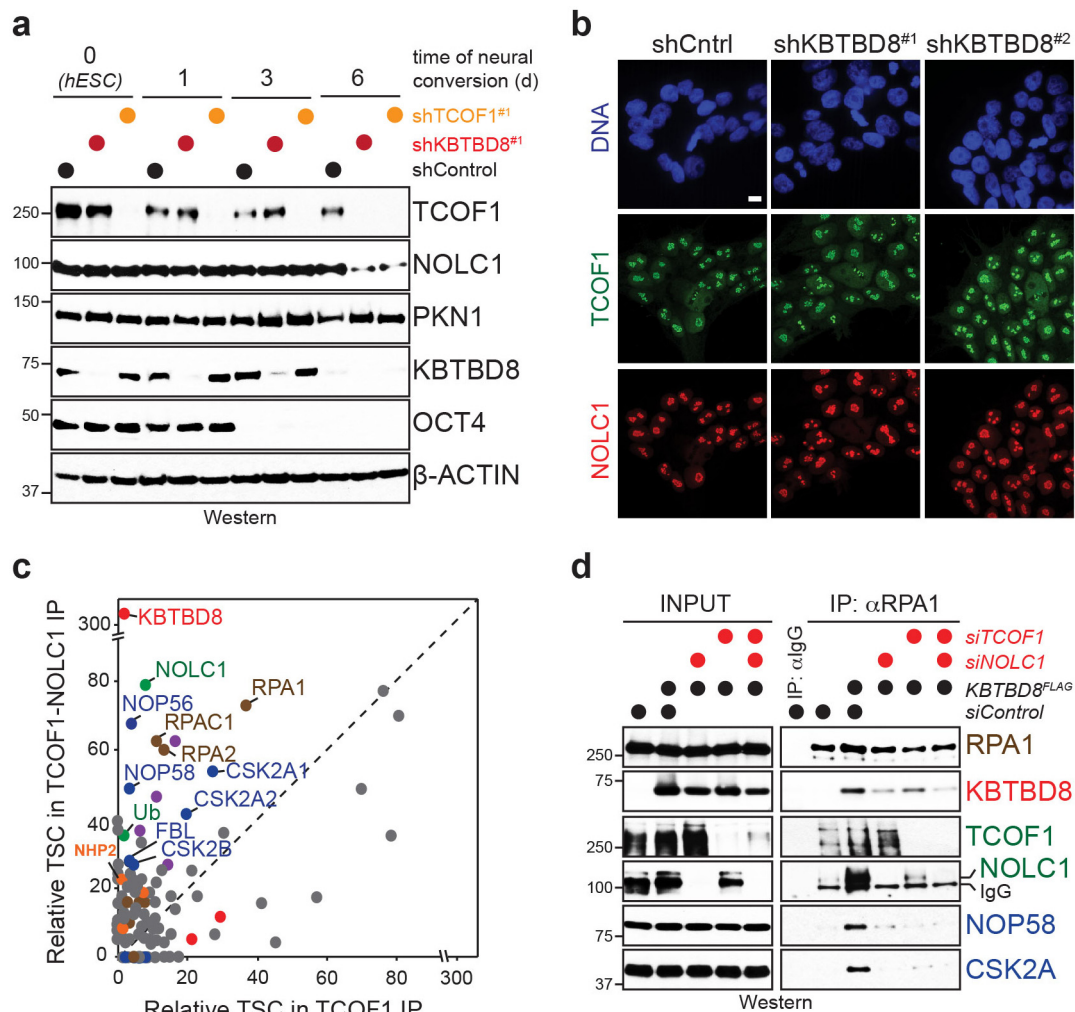
Extended Data Figure 5 | Biochemical characterization of the substrate adaptor role of KBTBD8. **a**, Domain structure of KBTBD8, including the residues mutated to generate ubiquitylation- (Y74A) and substrate-binding-deficient KBTBD8 (F550A, W579A). **b**, Effects of point mutations in predicted KELCH domain loops on binding of KBTBD8 to candidate substrates were determined by affinity purification and western blot analysis. **c**, Effects of point mutations in BTB domain on binding of KBTBD8 to CUL3 were determined by affinity purification and western blotting. Dimerization of Flag-KBTBD8 with KBTBD8-HA was analysed in the same experiment to provide a folding control. **d**, Binding of recombinant CUL3 to immobilized recombinant MBP-KBTBD8 variants was analysed by Coomassie. **e**, Binding of *in vitro*-transcribed/translated ³⁵S-NOLC1 to immobilized recombinant

KBTBD8 variants was analysed by autoradiography. **f**, Binding of *in vitro*-transcribed/translated ³⁵S-TCOF1 to immobilized recombinant KBTBD8 variants was analysed by autoradiography. **g**, Endogenous β-arrestin proteins in reticulocyte lysates binds immobilized, recombinant KBTBD8, as detected by western blot analysis. **h**, 293T cells were transfected with control- or β-arrestin 1/2-siRNAs and reconstituted with Flag-KBTBD8. Binding of KBTBD8 to endogenous TCOF1 and NOLC1 was analysed by anti-Flag affinity purification and western blot analysis. **i**, Ubiquitylation of HA-TCOF1 in 293T cells depleted of β-arrestin 1/2 and reconstituted with KBTBD8 was determined after denaturing Ni-NTA purification by western blotting as described above. **j**, Ubiquitylation of HA-NOLC1 was detected in 293T cells depleted of β-arrestins and reconstituted with KBTBD8, as described above.



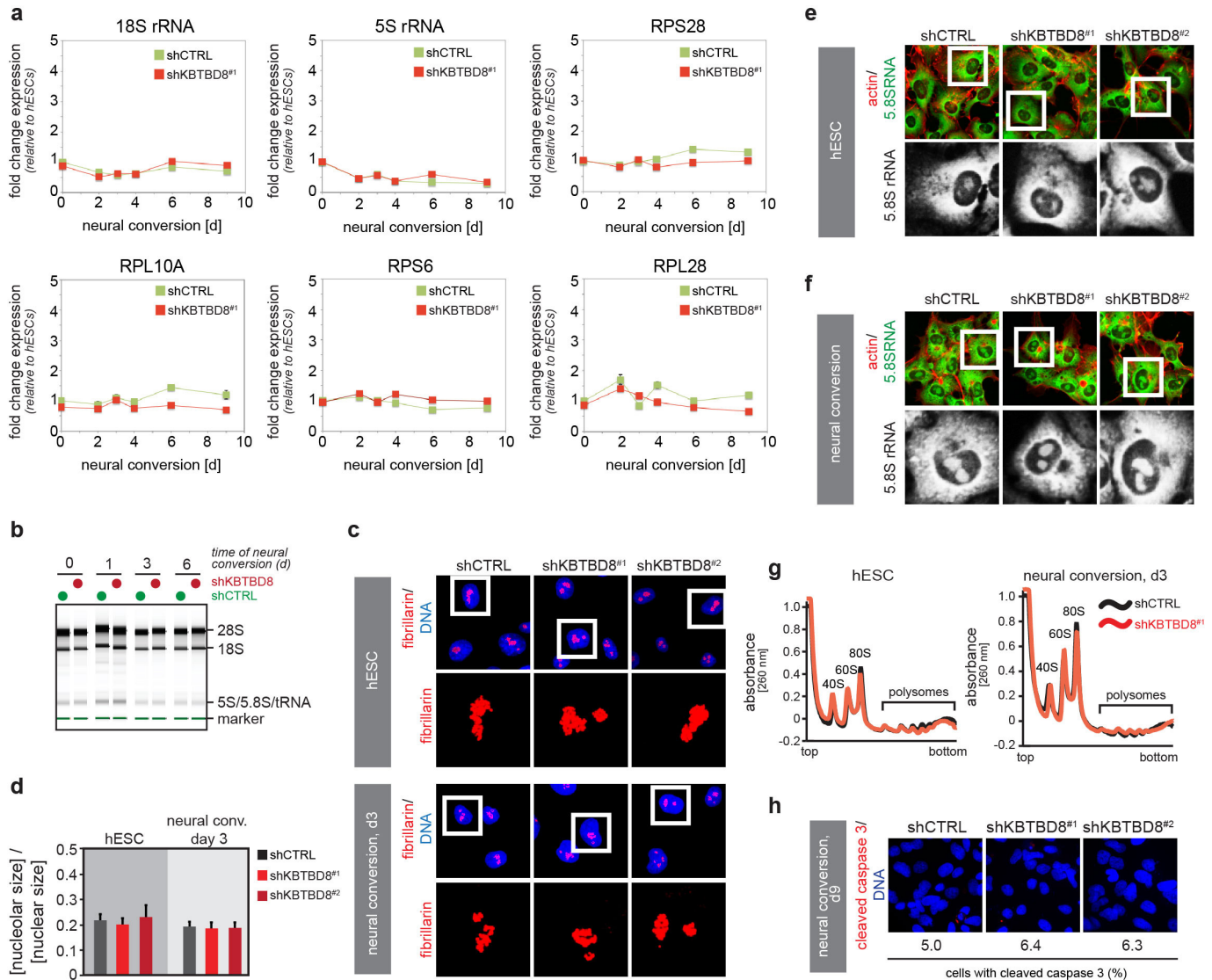
Extended Data Figure 6 | KBTBD8 specifies neural crest fate through TCOF1 and NOLC1. **a**, mRNA levels of *KBTBD8*, *NOLC1* and *TCOF1* were determined in hESCs or differentiating cells transduced with lentiviruses expressing the indicated shRNAs by qRT-PCR (mean of 3 technical replicates \pm s.e.m.). **b**, hESCs stably depleted of *KBTBD8* and reconstituted with either wild-type *KBTBD8*, *KBTBD8*(W579A), or *KBTBD8*(Y74A) were subjected to neural conversion (9 days) and analysed for the expression of marker proteins by qRT-PCR (mean of 3 technical replicates \pm s.e.m.). **c**, hESCs stably depleted of *KBTBD8*, *TCOF1*, or *NOLC1* were subjected to neural conversion (9 days) and analysed for marker expression by qRT-PCR

(mean of 3 technical replicates \pm s.e.m.). **d**, Depletion of *TCOF1* or *NOLC1* from hESCs results in loss of neural crest cells, as determined by triple staining immunofluorescence against the neural crest markers HNK1, TFAP2 and p75 ($n > 200$ cells, mean of 3 biological replicates \pm s.d.). Scale bar, 10 μ m. **e**, hESCs were transduced with lentiviruses expressing control or *BRD2* shRNAs, subjected to puromycin selection for 7 days, and analysed by western blot analysis. **f**, Depletion efficiency for shRNAs against various *KBTBD8* binding partners, as determined by qRT-PCR (mean of 3 technical replicates \pm s.e.m.).



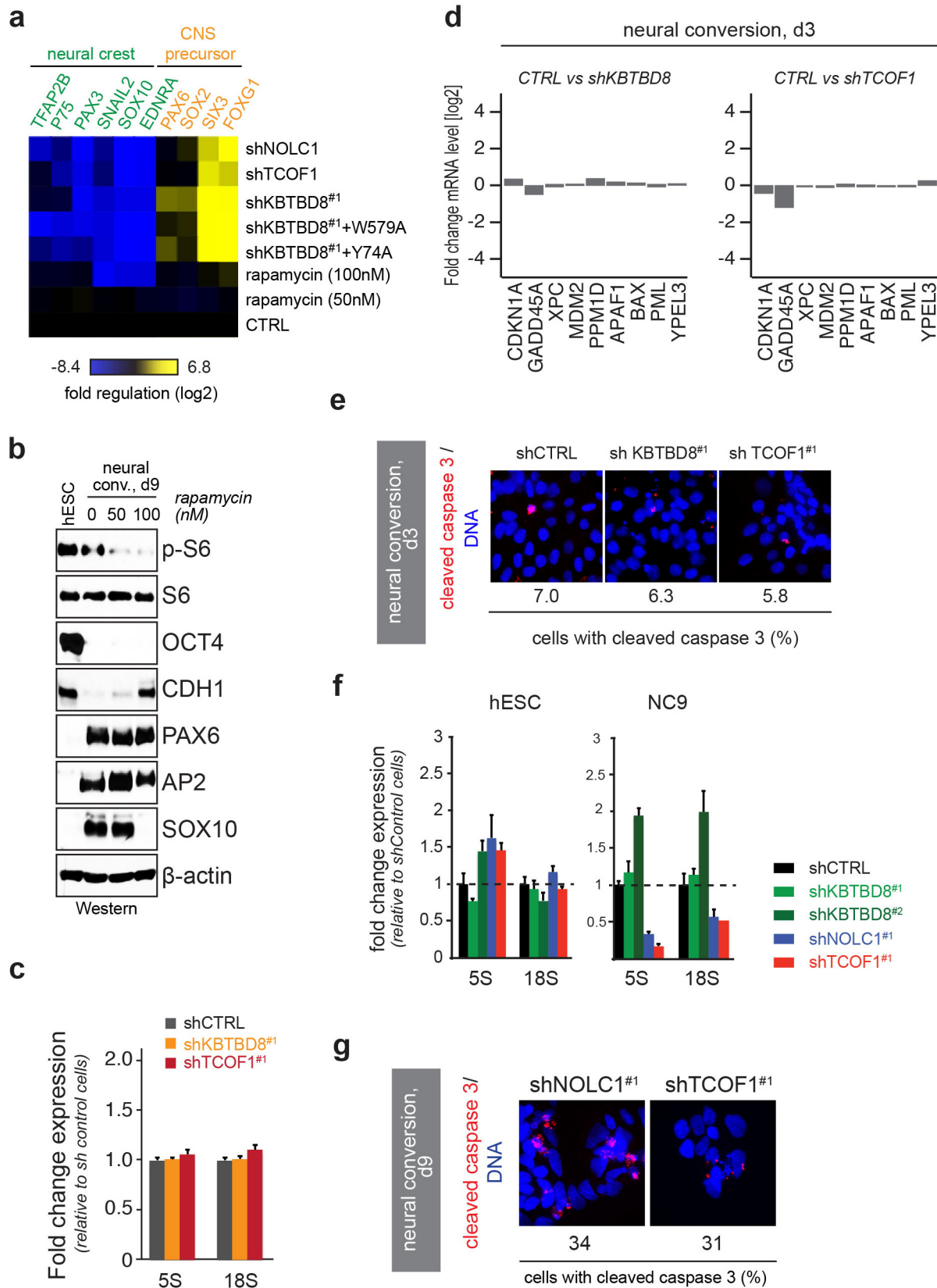
Extended Data Figure 7 | Characterization of TCOF1 regulation by ubiquitylation. **a**, hESCs depleted of either KBTBD8 or TCOF1 were subjected to neural conversion and analysed for expression of indicated proteins by western blot. **b**, Control or KBTBD8-depleted hESCs were fixed and subjected to indirect immunofluorescence analysis against endogenous TCOF1 or NOLC1. Scale bar, 10 μ m. **c**, Total spectral counts of proteins associated with TCOF1–NOLC1 complexes purified by sequential immunoprecipitation in the

presence of KBTBD8 compared to single TCOF1 affinity purification in the absence of KBTBD8, as determined by mass spectrometry (sum of 3 biological replicates). **d**, 293T cells were reconstituted with CUL3^{KBTBD8} and depleted of TCOF1 and NOLC1 by siRNAs. Endogenous RNA polymerase I was immunoprecipitated and binding to the SSU processome (NOP58, CSK2A) was analysed by western blot.



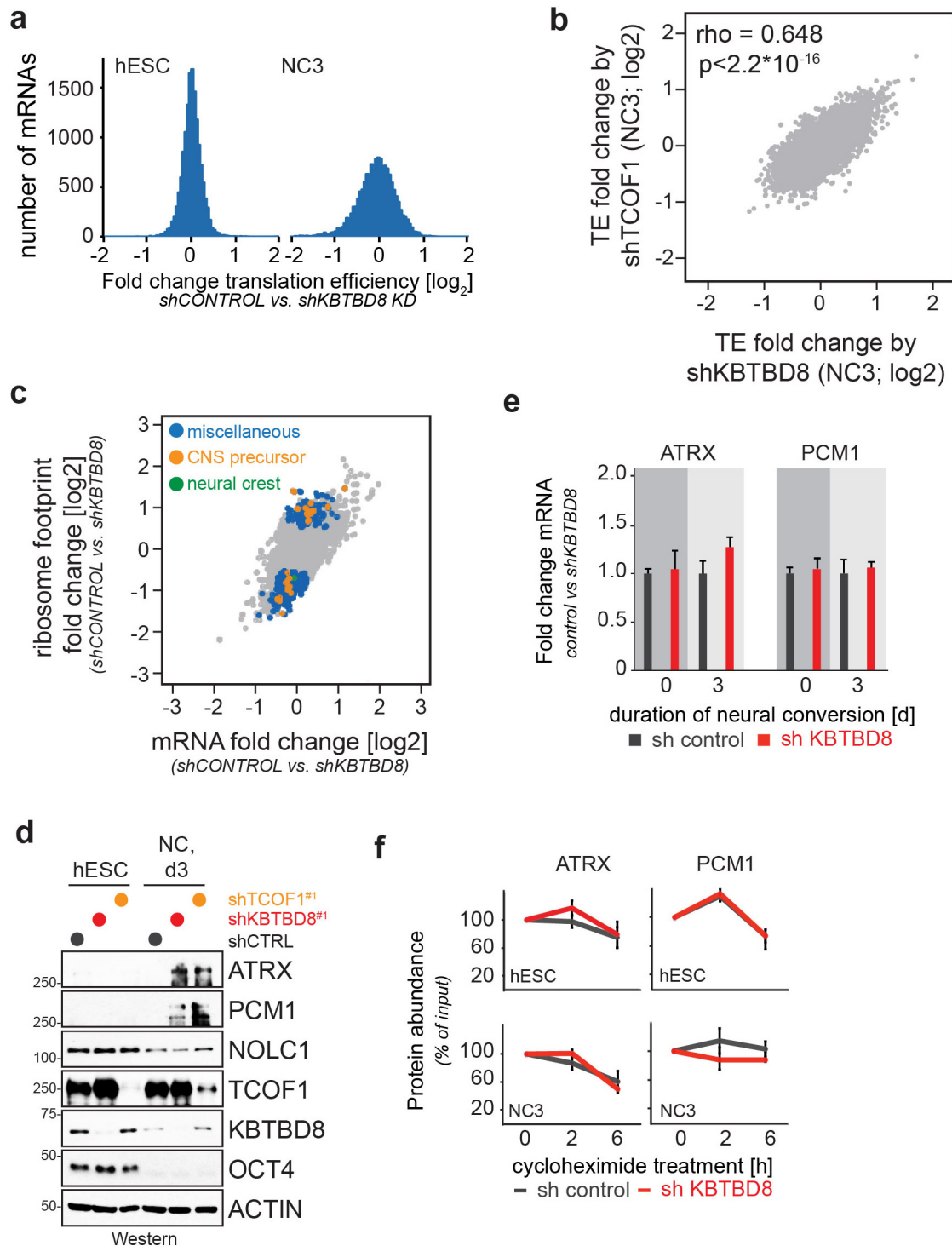
Extended Data Figure 8 | KBTBD8 is not required for general ribosome biogenesis. **a**, hESCs stably depleted of KBTBD8 were subjected to neural conversion and levels of 5S rRNA, 18S rRNA and mRNAs encoding RPS6, RPS28, RPL10A and RPL28 were measured by qRT-PCR (mean of 3 technical replicates \pm s.e.m.). **b**, hESCs stably depleted of KBTBD8 were subjected to neural conversion, and total RNA was subjected to a bioanalyser assay to monitor processing of ribosomal RNAs. **c**, hESCs stably depleted of KBTBD8 were subjected to neural conversion (3 days), and nucleoli were analysed by anti-fibrillar (original magnification: 60 \times , confocal) immunofluorescence microscopy. **d**, Quantification of nucleolar analysis described above (mean of 3 technical replicates \pm s.e.m.). **e**, hESCs stably depleted of KBTBD8 were analysed for localization of 5.8S rRNA by anti-5.8S rRNA immunofluorescence

microscopy (original magnification: 60 \times , confocal). **f**, hESCs depleted of KBTBD8 were subjected to neural conversion and analysed by anti-5.8S rRNA immunofluorescence microscopy (original magnification: 60 \times , confocal). **g**, Polysomes were purified from control or KBTBD8-depleted hESCs and differentiated counterparts subjected to neural conversion via sucrose gradient centrifugation followed by fractionation and UV detection. **h**, KBTBD8-depleted hESCs were subjected to neural conversion for 9 days and analysed for apoptosis by immunofluorescence analysis against cleaved caspase 3 (red) and DNA (Hoechst, blue). Cells with active caspase 3 staining were quantified (~200 cells per condition; scale bar, 10 μ m) (original magnification: 40 \times , confocal).



Extended Data Figure 9 | Characterization of KBTBD8- and TCOF1-depleted hESCs during neural conversion. **a**, hESCs were treated with increasing concentrations of rapamycin, subjected to neural conversion for 9 days, and analysed for expression of neural crest or CNS precursor markers by qRT-PCR. For comparison, effects of KBTBD8, TCOF1, or NOLC1 depletion (extracted from Fig. 3a) are shown. **b**, hESCs were treated with rapamycin, subjected to neural conversion, and analysed for marker expression by western blotting. **c**, hESCs were depleted of KBTBD8 or TCOF1, subjected to neural conversion for 3 days, and analysed for expression of 5S and 18S rRNA by qRT-PCR (mean of 3 technical replicates \pm s.e.m.). **d**, hESCs depleted of KBTBD8 or TCOF1 were subjected to neural conversion for 3 days and

analysed for p53 activation by RNA-seq against p53 targets. **e**, hESCs were depleted of KBTBD8 or TCOF1, subjected to neural conversion for 3 days, and analysed for apoptosis by immunofluorescence microscopy against cleaved caspase 3. Quantification shown below (~ 200 cells per condition). **f**, hESCs depleted of KBTBD8 were subjected to neural conversion for 9 days and analysed for expression levels of 5S and 18S rRNA by qRT-PCR (mean of 3 technical replicates \pm s.e.m.). **g**, hESCs stably depleted of NOLC1 or TCOF1 were subjected to neural conversion for 9 days and analysed by immunofluorescence microscopy against cleaved caspase 3 (red) or DNA (Hoechst, blue). Quantification is shown below (~ 200 cells per condition; scale bar, 10 μ m).



Extended Data Figure 10 | KBTBD8 controls translation. **a**, hESCs stably depleted of KBTBD8 were subjected to neural conversion for 3 days, and hESCs and differentiating cells were analysed by RNA deep sequencing and ribosomal profiling to determine translation efficiency. Distribution of translation efficiency changes for 7,725 mRNAs brought about by KBTBD8 depletion is shown. **b**, hESCs stably depleted of either TCOF1 or KBTBD8 were subjected to neural conversion for 3 days, and translation efficiency was determined by RNA-seq and ribosome profiling. **c**, Translation efficiency blot of differentiating hESCs transduced with control or *KBTBD8* shRNAs was labelled for significantly affected transcripts in general (blue), with links to CNS precursor formation (gold), or with links to neural crest formation (green).

d, hESCs stably depleted of KBTBD8 or TCOF1 were subjected to neural conversion for 3 days, and expression levels of indicated proteins were analysed by western blotting. **e**, hESCs stably depleted of KBTBD8 were subjected to neural conversion for 3 days, and levels of *ATRAX1* and *PCM1* mRNA were determined by qRT-PCR (mean of 3 technical replicates \pm s.e.m.). **f**, hESCs stably depleted of KBTBD8 were subjected to neural conversion for 3 days, and protein stability of *ATRAX1* and *PCM1* was determined by cycloheximide chase and western blotting (mean of 3 biological replicates \pm s.d., *ATRAX1* and *PCM1* levels were normalized relative to actin levels and 0 h time point set to 100%).

Neutrophil ageing is regulated by the microbiome

Dachuan Zhang^{1,2}, Grace Chen^{1,2}, Deepa Manwani³, Arthur Mortha^{4,5}, Chunliang Xu^{1,2}, Jeremiah J. Faith^{5,6}, Robert D. Burk³, Yuya Kunisaki^{1,2,†}, Jung-Eun Jang^{1,2}, Christoph Scheiermann^{1,2,†}, Miriam Merad^{4,5} & Paul S. Frenette^{1,2,7}

Blood polymorphonuclear neutrophils provide immune protection against pathogens, but may also promote tissue injury in inflammatory diseases^{1,2}. Although neutrophils are generally considered to be a relatively homogeneous population, evidence for heterogeneity is emerging^{3,4}. Under steady-state conditions, neutrophil heterogeneity may arise from ageing and replenishment by newly released neutrophils from the bone marrow⁵. Aged neutrophils upregulate CXCR4, a receptor allowing their clearance in the bone marrow^{6,7}, with feedback inhibition of neutrophil production via the IL-17/G-CSF axis⁸, and rhythmic modulation of the haematopoietic stem-cell niche⁵. The aged subset also expresses low levels of L-selectin^{5,9}. Previous studies have suggested that *in vitro*-aged neutrophils exhibit impaired migration and reduced pro-inflammatory properties^{6,10}. Here, using *in vivo* ageing analyses in mice, we show that neutrophil pro-inflammatory activity correlates positively with their ageing whilst in circulation. Aged neutrophils represent an overly active subset exhibiting enhanced $\alpha_M\beta_2$ integrin activation and neutrophil extracellular trap formation under inflammatory conditions. Neutrophil ageing is driven by the microbiota via Toll-like receptor and myeloid differentiation factor 88-mediated signalling pathways. Depletion of the microbiota significantly reduces the number of circulating aged neutrophils and dramatically improves the pathogenesis and inflammation-related organ damage in models of sickle-cell disease or endotoxin-induced septic shock. These results identify a role for the microbiota in regulating a disease-promoting neutrophil subset.

Neutrophils are a critical component of innate immunity. However, activated neutrophils can also promote certain diseases by secreting pro-inflammatory cytokines, and by interacting with other immune or blood cells². For example, activation of $\alpha_M\beta_2$ (Mac-1) integrin enables adherent neutrophils to interact with platelets and red blood cells (RBCs)¹¹. In sickle-cell disease (SCD), a severe blood disorder originating from a single mutation in the β -globin gene (*Hbb*)¹², the capture of sickle RBCs by activated Mac-1 on adherent neutrophils leads to acute vaso-occlusion, resulting in life-threatening crises^{11,13,14}. Intravital microscopy analyses have revealed considerable heterogeneity in Mac-1 activation on neutrophils recruited to the same venules of SCD mice, suggesting that subsets of neutrophils differ markedly in their pro-inflammatory activity¹¹.

To investigate whether the heterogeneity in pro-inflammatory activity of neutrophils is associated with their ageing, we first validated that neutrophils progressively lost L-selectin (CD62L) expression and upregulated CXCR4 as they aged whilst in circulation⁵ (Extended Data Fig. 1a, b). We analysed CD62L^{lo} neutrophils *in vivo* using multi-channel fluorescence intravital microscopy (MFIM) analysis^{11,13}. Notably, CD62L expression on adherent neutrophils in tumour necrosis factor α (TNF α)-inflamed post-capillary venules inversely correlated with Mac-1 activation (Fig. 1a, b), as determined using fluorescent microsphere beads that specifically bound to activated

Mac-1¹¹. In addition, a similar inverse correlation was observed in the ability of adherent neutrophils to capture RBCs (Fig. 1b).

Next, we analysed neutrophil populations in mice lacking P-selectin (*Selp*^{-/-}), an adhesion molecule essential for neutrophil recruitment under steady-state conditions^{5,15}. We found that the CD62L^{lo}CXCR4^{hi} aged neutrophil population was dramatically expanded in *Selp*^{-/-} mice, and neutrophils harvested from *Selp*^{-/-} mice showed significantly higher Mac-1 activity compared to those harvested from wild-type mice (Fig. 1c and Extended Data Fig. 1c, d). In addition, we analysed neutrophil populations after depletion of macrophages—which mediate neutrophil clearance¹⁶—using animals expressing a knocked-in diphtheria toxin receptor at the *Cd169* locus (*CD169-DTR*)¹⁷. We found that the aged neutrophil population was significantly expanded without elevation of major inflammatory cytokines, and Mac-1 activation was significantly increased on adherent neutrophils in macrophage-depleted mice (Fig. 1d and Extended Data Fig. 1e, f). These data suggest that CD62L^{lo}CXCR4^{hi} aged neutrophils exhibit enhanced Mac-1 activation during inflammation.

To evaluate the specificities of ageing versus the activation of an inflammatory program, we compared the transcriptome of control, aged and TNF α -activated neutrophils. We transfused whole blood and collected donor neutrophils 6 h later to derive *in vivo*-aged neutrophils, and compared them to control neutrophils that were transferred for only 10 min. Additionally, we obtained neutrophils from TNF α -treated mice for comparison with neutrophils activated by systemic inflammation. Gene set enrichment analyses¹⁸ revealed that aged neutrophils differed from activated neutrophils in many aspects, such as cytokine and chemokine secretion, Ras and P38/MAPK signalling pathways (Fig. 1e). However, aged neutrophils upregulated several pathways that were also enhanced during neutrophil activation, including integrin and leukocyte adhesion, Toll-like receptor (TLR) and NOD-like receptor (NLR), and NF κ B signalling pathways (Fig. 1e, f and Extended Data Table 1, 2). Analysis of surface antigens revealed that aged neutrophils exhibited significantly higher levels of TLR4 and molecules involved in cell migration and intercellular interactions, including CD11b, CD49d, and Icam1 (Fig. 1g and Extended Data Fig. 1g–i). These results demonstrate that neutrophils constitutively receive priming signals and become more active as they age in circulation.

The upregulation of several inflammatory pathways in aged neutrophils suggests a contribution by exogenous inflammatory mediators. Microbiota-derived molecules may cross the intestinal barrier to exert systemic influences, affecting multiple immune populations including T cells, innate lymphoid cells and macrophages^{19,20}. Recent studies suggest that neutrophil production and the phagocytic capacity of bone-marrow-derived neutrophils may be regulated by the microbiota^{21–24}, raising the possibility that these factors also influence the ageing process of circulating neutrophils.

We sought to test this hypothesis by treating mice with broad-spectrum antibiotics (ABX) for 4–6 weeks^{19,23}, which led to highly

¹Ruth L. and David S. Gottesman Institute for Stem Cell and Regenerative Medicine Research, Albert Einstein College of Medicine, Bronx, New York 10461, USA. ²Department of Cell Biology, Albert Einstein College of Medicine, Bronx, New York 10461, USA. ³Department of Pediatrics, Albert Einstein College of Medicine, Bronx, New York 10461, USA. ⁴Department of Oncological Sciences, Mount Sinai School of Medicine, New York, New York 10029, USA. ⁵The Immunology Institute, Mount Sinai School of Medicine, New York, New York 10029, USA. ⁶The Institute for Genomics and Multiscale Biology, Mount Sinai School of Medicine, New York, New York 10029, USA. ⁷Department of Medicine, Albert Einstein College of Medicine, Bronx, New York 10461, USA. [†]Present addresses: Department of Medicine and Biosystemic Science, Kyushu University, Fukuoka, Fukuoka 812-8582, Japan (Y.K.); Walter Brendel Centre of Experimental Medicine, Ludwig-Maximilians-University, 81377 Munich, Germany (C.S.).

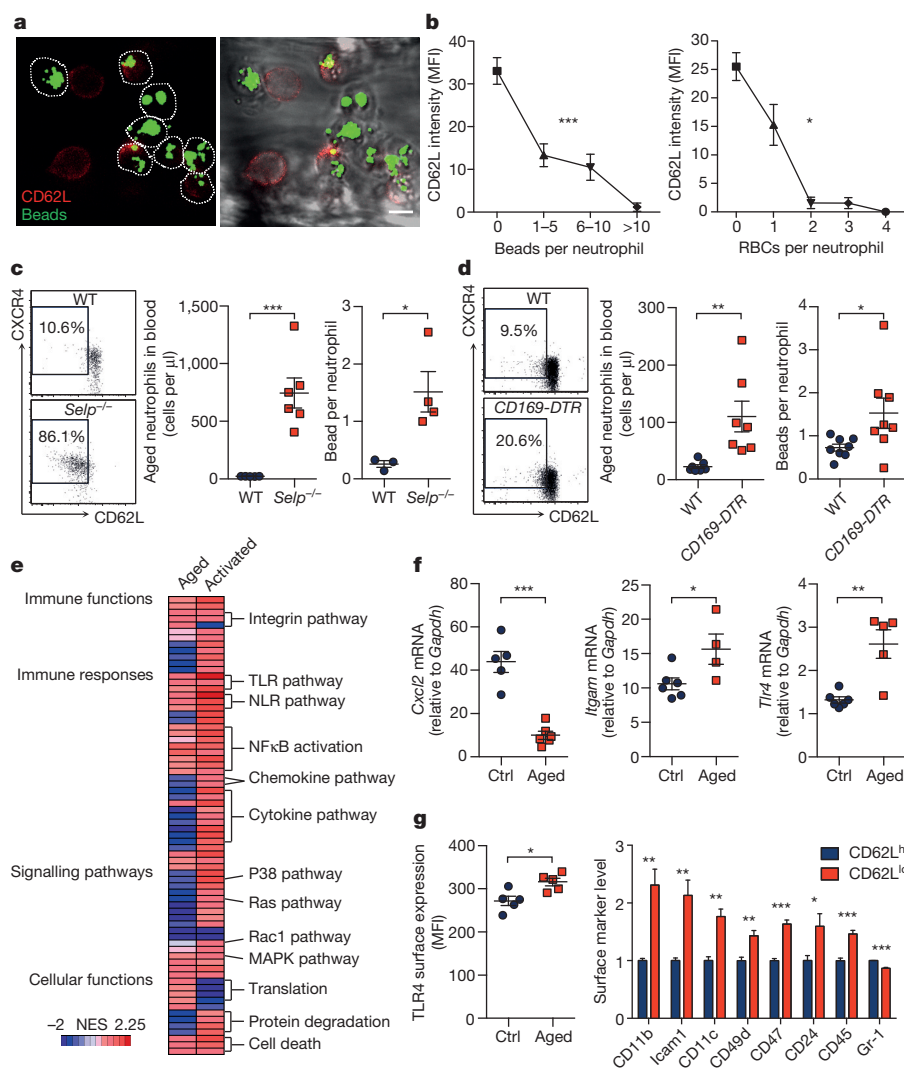


Figure 1 | Aged neutrophils represent an overly active subset of neutrophils.

a, MFIM analysis of CD62L expression (red) and Mac-1 specific albumin-coated fluorescent microsphere beads (green) captured by adherent neutrophils (dashed lines). Left, fluorescence channel; right, fluorescence combined with brightfield channels. MFI, mean fluorescence intensity. Scale bar, 10 μ m.

b, Correlation between CD62L expression and bead capture or neutrophil-RBC interaction ($n = 126$ cells from 3 mice).

c, d, Flow cytometry analysis of CD62L^{lo}CXCR4^{hi} aged neutrophils and MFIM analysis of Mac-1 activation on neutrophils from wild-type (WT) and *Selp*^{-/-} mice (**c**; middle, $n = 5, 6$ mice, respectively; right, $n = 3, 4$ mice, respectively), or in diphtheria-toxin-treated wild-type and *CD169-DTR* mice (**d**; middle, $n = 7$ mice; right, $n = 8$ mice).

efficient depletion and dramatic alterations in the composition of the gut microbiota (Extended Data Fig. 2a–d). Microbiota depletion resulted in significant and selective reductions of neutrophil numbers in circulation and bone marrow, and a significant reduction in spleen cellularity with decreased numbers of multiple leukocyte populations (Extended Data Fig. 3a–d). Notably, both the percentages and numbers of aged neutrophils were significantly reduced in ABX-treated mice, and the numbers were completely restored when the TLR4 ligand lipopolysaccharide (LPS) was added back by intragastric gavage (Fig. 2a). We further analysed neutrophil-LPS interaction by administering fluorescently labelled LPS, and found that as soon as 1 h after LPS gavage, specific fluorescence signals were detectable on neutrophils in circulation, the spleen and bone marrow (Extended Data Fig. 3e). In addition, we found that the reintroduction of the TLR2 ligand peptidoglycan, but not the NOD1/2 activator mTriDAP, could also restore the numbers of aged neutrophils in ABX-treated mice

e, Heat map of normalized enrichment scores (NES) for selected pathways in aged and TNF α -activated neutrophils, as compared to control neutrophils ($n = 3$ mice). Red, upregulation; blue, downregulation. **f, g**, *Cxcl2*, *Itgam* and *Tlr4* mRNA expression levels determined by quantitative PCR in control and aged neutrophils (**f**; left, $n = 5, 6$ mice, respectively; middle, $n = 6, 4$ mice, respectively; right, $n = 6, 5$ mice, respectively), and surface expression levels of TLR4 and selected adhesion molecules determined by flow cytometry on CD62L^{lo} aged and CD62L^{hi} young neutrophils (**g**; left, $n = 5$ mice; right, $n = 4$ mice). Error bars, mean \pm s.e.m. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$, data representing two or more independent experiments analysed with one-way ANOVA (**b**) or unpaired Student's *t*-test (**c, d, f, g**).

(Extended Data Fig. 3f), suggesting that multiple microbiota-derived molecules may contribute to neutrophil ageing under steady-state conditions.

To validate that the microbiota could indeed regulate neutrophil ageing, we analysed neutrophil populations in germ-free mice. Compared to specific-pathogen-free animals, germ-free mice exhibited broad alterations in both innate and adaptive immune cells (Extended Data Fig. 4a–c). Consistently, the numbers of total and aged neutrophils were significantly reduced in germ-free mice, and the numbers were partially restored when germ-free mice were reconstituted by fecal transplantation. In addition, treating germ-free mice with ABX did not further reduce their aged neutrophil numbers (Fig. 2b and Extended Data Fig. 4d, e). Furthermore, we transfused whole blood obtained from wild-type donor mice into wild-type, ABX-treated or germ-free recipients. The percentages of chronologically aged donor neutrophils progressively increased after the transfusion,

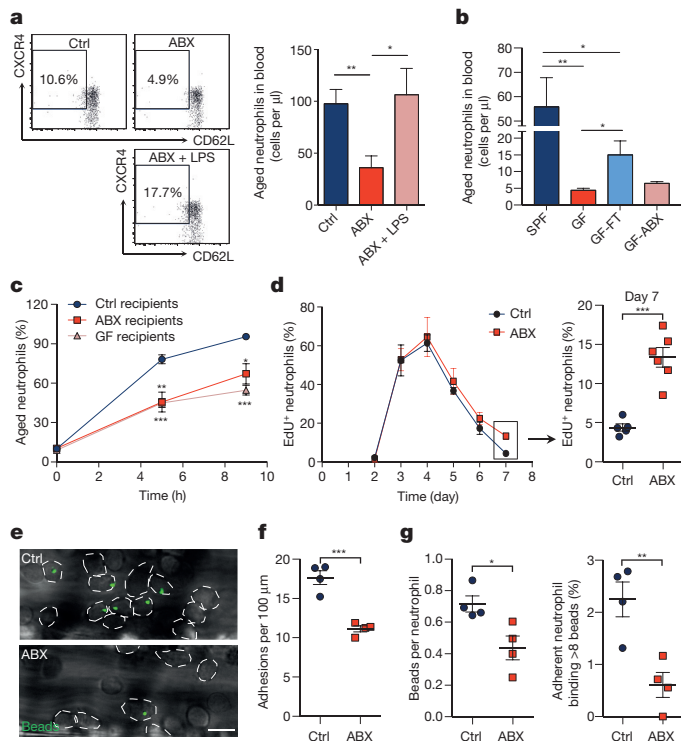


Figure 2 | Neutrophil ageing is driven by the microbiota. **a**, Flow cytometry analysis of aged neutrophils in control, antibiotics (ABX)-treated mice and ABX-treated mice fed with LPS ($n = 12, 7, 5$ mice, respectively). **b**, Numbers of aged neutrophils in specific-pathogen-free (SPF) mice, germ-free (GF) mice, germ-free mice reconstituted by fecal transplantation (GF-FT), and germ-free mice treated with antibiotics (GF-ABX) ($n = 5, 5, 5, 3$ mice, respectively). **c**, Ageing kinetics of donor neutrophils after adoptive transfer into control, ABX-treated or germ-free recipients ($n = 4$ mice). **d**, EdU pulse-chase analysis of neutrophil release–clearance kinetics in control and ABX-treated mice (Ctrl, $n = 5, 5, 6, 8, 5$ mice; ABX, $n = 5, 5, 5, 6, 6$ mice for days 3, 4, 5, 6 and 7, respectively). **e–g**, Representative images (**e**) and quantification of neutrophil adhesion (dotted lines; **f**) and Mac-1 activation on adherent neutrophils (**g**) in control and ABX-treated mice ($n = 4$ mice). Scale bar, 10 μm . Error bars, mean \pm s.e.m. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$, data representing two or more independent experiments analysed with unpaired Student's *t*-test.

but at a significantly slower rate in ABX-treated and germ-free recipients (Fig. 2c). These results strongly suggest that neutrophil ageing is delayed in a bacterially depleted environment.

Since alterations in neutrophil clearance may influence the number of aged neutrophils (Fig. 1c, d), we investigated whether ABX treatment modulates neutrophil ageing by acting on clearance mechanisms. We first analysed adhesion molecule expression on endothelial cells and observed no difference between control and ABX-treated mice (Extended Data Fig. 5a). Next, we analysed macrophage numbers in the spleen, bone marrow and liver, the organs that clear neutrophils²⁵. We found that macrophage numbers decreased by $\sim 36\%$ in the spleen, increased by $\sim 30\%$ in the bone marrow, and did not change in the liver of ABX-treated mice (Extended Data Fig. 5b, c). Furthermore, we depleted macrophages in ABX-treated mice using the *CD169-DTR* model, and found that ABX treatment significantly reduced aged neutrophil numbers in macrophage-depleted animals (Extended Data Fig. 5c, d), suggesting that microbiota-driven ageing and macrophage-mediated clearance are independent mechanisms that regulate the number of aged neutrophils. We then analysed the release–clearance kinetics of circulating neutrophils using a 5-ethynyl-2'-deoxyuridine (EdU) pulse-chase labelling strategy¹⁶. We observed significantly more EdU⁺ neutrophils remaining in circulation on day 7, suggesting a delayed clearance in ABX-treated mice (Fig. 2d). In

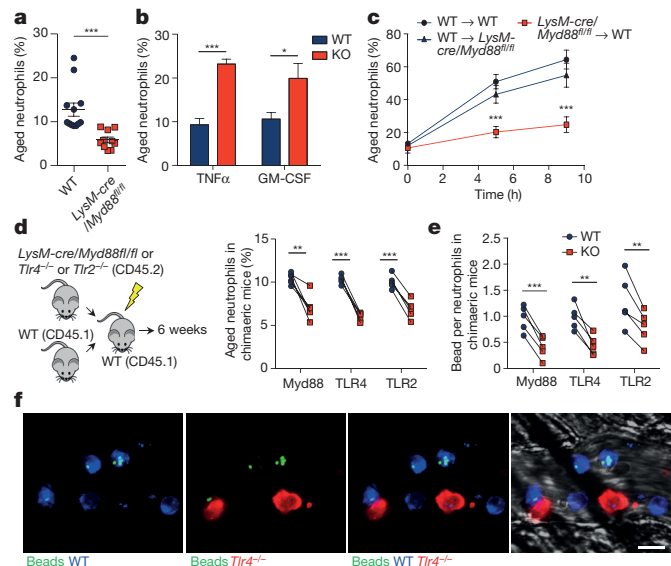


Figure 3 | Microbiota-driven neutrophil ageing is mediated by neutrophil TLRs and Myd88 signalling. **a**, Percentages of aged neutrophils in wild-type and *LysM-cre/Myd88^{fl/fl}* mice, as analysed by flow cytometry ($n = 12, 10$ mice, respectively). **b**, Percentages of aged neutrophils in wild-type and *Tnf^{-/-}* or *Csf2^{-/-}* mice ($n = 5$ mice). **c**, Ageing kinetics of donor neutrophils after adoptive transfer from either wild-type or *LysM-cre/Myd88^{fl/fl}* mice into wild-type or *LysM-cre/Myd88^{fl/fl}* recipients ($n = 6$ mice). **d**, Percentages of the aged subset in wild-type and *LysM-cre/Myd88^{fl/fl}*, *Tlr4^{-/-}* or *Tlr2^{-/-}* neutrophils in chimaeric mice ($n = 5$ mice). **e**, MFIM analysis of Mac-1 activation on wild-type and *LysM-cre/Myd88^{fl/fl}*, *Tlr4^{-/-}* or *Tlr2^{-/-}* neutrophils in chimaeric mice ($n = 5$ mice). **f**, Representative images showing wild-type (CD45.1⁺, blue) and *Tlr4^{-/-}* (CD45.2⁺, red) neutrophils and beads (green). Scale bar, 10 μm . Error bars, mean \pm s.e.m. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$, data representing two or more independent experiments analysed with unpaired Student's *t*-test (**a–c**) or paired Student's *t*-test (**d, e**).

addition, we investigated the functional impact of microbiota depletion using intravital microscopy, and observed significant reductions in neutrophil adhesion and Mac-1 activation in ABX-treated compared to control mice (Fig. 2e–g). These data suggest that neutrophil ageing, which leads to the generation of a functionally overly active subset of neutrophils, is driven by the microbiota.

Neutrophils express multiple pattern recognition receptors, including TLR2 and TLR4²⁶, which may directly transduce microbiota-derived signals. Alternatively, microbiota-derived signals may stimulate certain immune cells to secrete pro-inflammatory cytokines such as TNF α and granulocyte-macrophage colony-stimulating factor (GM-CSF)^{19,27}, which could in turn prime circulating neutrophils. To investigate how microbiota-derived signals regulate neutrophil ageing, we characterized aged neutrophils in *LysM-cre/Myd88^{fl/fl}* mice, in which Myd88, a signalling molecule that mediates most TLR signalling, is specifically deleted in myeloid cells. Notably, we observed significant reductions in the percentages and numbers of aged neutrophils in these mice (Fig. 3a and Extended Data Fig. 5e). Similarly, we also found significant reductions of aged neutrophils in TLR2- and TLR4-knockout mice (Extended Data Fig. 5f). By contrast, the aged neutrophil population was expanded in TNF α - and GM-CSF-knockout mice (Fig. 3b and Extended Data Fig. 5g), suggesting that the microbiota may not drive neutrophil ageing by stimulating TNF α or GM-CSF secretion.

To further delineate how Myd88 mediates microbiota-driven ageing, we adoptively transferred either *LysM-cre/Myd88^{fl/fl}* or wild-type neutrophils into wild-type or *LysM-cre/Myd88^{fl/fl}* recipient mice, and analysed donor neutrophil ageing *in vivo*. Interestingly, neutrophil ageing was almost completely abrogated when Myd88-deficient neutrophils were transferred into wild-type recipients, whereas the ageing kinetics remained largely unaffected when wild-type neutrophils were

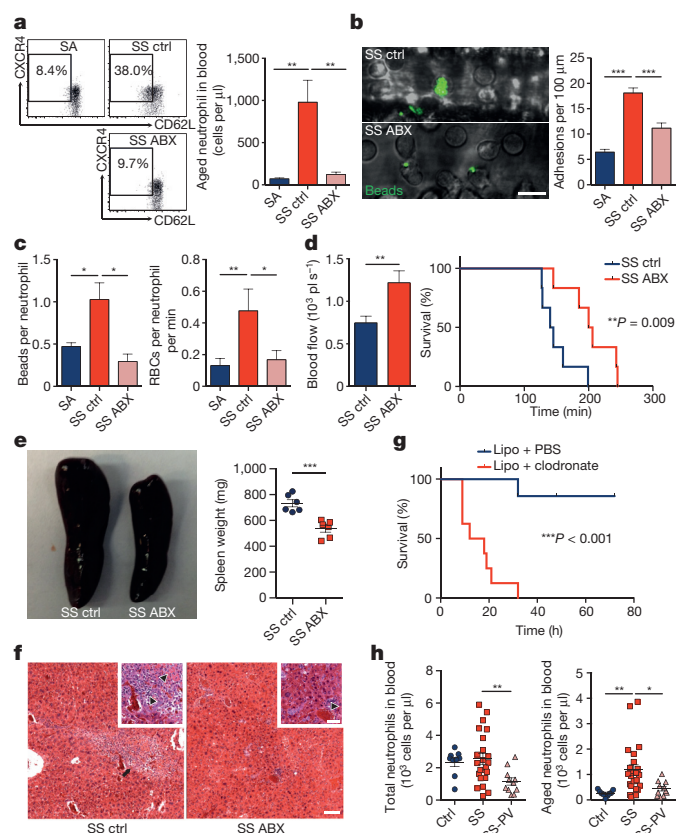


Figure 4 | Microbiota depletion reduces vaso-occlusive events in sickle-cell disease. **a**, Flow cytometry analysis of aged neutrophils in hemizygous control (SA), control SCD (SS ctrl) and antibiotic-treated SCD (SS ABX) mice ($n = 8, 9, 9$ mice, respectively). **b**, MFIM analysis of neutrophil adhesion and Mac-1 activation on adherent neutrophils in SA, SS control and SS ABX mice ($n = 6, 8, 10$ mice, respectively). Scale bar, 10 μm . **c**, Mac-1 activation on adherent neutrophils and neutrophil–RBC interaction in SA, SS control and SS ABX mice (left, $n = 4, 3, 3$ mice, respectively; right, $n = 69, 42, 54$ vessels from 6, 7 and 7 mice, respectively). **d**, Blood flow and survival time of SS control and SS ABX mice in acute vaso-occlusive crisis (left, $n = 36, 48$ vessels from 7 and 8 mice, respectively; right, $n = 6$ mice). **e**, Representative images and weights of spleen in SS control and SS ABX mice ($n = 6$ mice). **f**, Haematoxylin and eosin staining showing liver damage in SS control and SS ABX mice. Arrow, liver fibrosis; arrowheads, necrosis and inflammation. Scale bars, 50 μm . **g**, Survival time of SS mice treated with PBS- or clodronate-encapsulated liposome (lipo; $n = 7, 8$ mice, respectively). **h**, Numbers of total and aged neutrophils in healthy controls, SCD patients (SS), and SCD patients on penicillin V prophylaxis (SS-PV; $n = 9, 23, 11$ subjects, respectively). Error bars, mean \pm s.e.m. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$, data representing two or more independent experiments analysed with unpaired Student's *t*-test (**a–d** (left), **e**, **h**) or log-rank test (**d** (right), **g**).

transferred into *LysM-cre/Myd88^{fl/fl}* recipients (Fig. 3c). Next, we generated chimaeric mice reconstituted with a mixture of wild-type and Myd88-, TLR4- or TLR2-deficient bone marrow cells, which enabled us to compare wild-type and deficient neutrophils in the same mouse, thus avoiding potential differences caused by the environment. Consistently, we observed significantly lower percentages of the aged subset in Myd88-, TLR2- and TLR4-deficient neutrophils, compared to wild-type neutrophils in the same chimaeric mice (Fig. 3d). By contrast, the percentages of total neutrophils in Myd88-, TLR2- and TLR4-deficient leukocytes were unaltered or expanded (Extended Data Fig. 5h), suggesting a specific effect of TLR and Myd88 deficiency on ageing, but not the generation, of neutrophils. We also subjected these chimaeric mice to intravital microscopy, and found significantly lower Mac-1 activity on Myd88-, TLR4- and TLR2-deficient neutrophils (Fig. 3e, f and Extended Data Fig. 5i). These findings

strongly suggest that neutrophil TLRs and Myd88 signalling mediate microbiota-driven neutrophil ageing.

In addition to analysing Mac-1 activation, we investigated whether ageing affects the ability of neutrophils to form neutrophil extracellular traps (NETs) in response to pathological stimulation. We enriched aged neutrophils by injecting antibodies to block P- and E-selectins⁵, and found that neutrophils obtained from anti-selectin-treated mice exhibited significantly increased reactive oxygen species (ROS) production (Extended Data Fig. 6a, b). We induced NET formation by treating neutrophils with LPS *in vitro*²⁸, and quantified NETs based on the co-localization of DNA with citrullinated histone H3 (CitH3) and neutrophil elastase²⁹. We found that NET formation was significantly increased in neutrophils collected from anti-selectin-treated mice. By contrast, neutrophils isolated from ABX-treated mice exhibited a marked reduction in NET formation (Extended Data Fig. 6c, d).

To investigate whether microbiota depletion impacts NET formation *in vivo*, we challenged control and ABX-treated mice with a lethal dose of LPS to induce septic shock¹⁵, and injected fluorophore-conjugated antibodies to image NETs in the liver vasculature. Notably, we observed a significant reduction in the number of NETs formed in septic liver after microbiota depletion, and dramatic decreases in soluble NET biomarkers—plasma nucleosome and plasma DNA (Extended Data Fig. 7a, b). In addition, immunofluorescence analyses revealed that the septic liver vasculature contained numerous aggregated CitH3⁺ neutrophils, which was commonly found to be associated with fibrin deposition (Extended Data Fig. 7c, d). By contrast, CitH3⁺ neutrophils, neutrophil aggregates and fibrin deposition were markedly reduced in ABX-treated mice, leading to significantly prolonged survival of these mice (Extended Data Fig. 7e–g). Remarkably, this improvement in survival was abrogated by infusing aged neutrophils, but not by infusing the same number of young neutrophils back into ABX-treated mice (Extended Data Fig. 7g).

To assess further the role of neutrophil ageing in a disease model, we analysed mice with SCD, a disease characterized by recurrent episodes of vaso-occlusion in which neutrophils play a primary function^{11,14}. While SCD mice exhibited significant expansion of all major leukocyte subsets compared to hemizygous mice, ABX-mediated microbiota depletion led to a significant and selective reduction of neutrophils, but not other leukocyte populations (Extended Data Fig. 8a). Notably, the numbers of aged neutrophils were expanded more than tenfold in SCD mice, and the expansion was completely abrogated by microbiota depletion (Fig. 4a).

To investigate the functional impact of reduced aged neutrophil numbers in disease outcome, we challenged hemizygous, untreated and ABX-treated SCD mice with TNF α and evaluated the cremaster microcirculation by intravital microscopy^{11,14}. SCD mice exhibited significant increases in neutrophil adhesion, Mac-1 activation and heterotypic interactions with RBCs compared to hemizygous mice, all of which were markedly reduced by microbiota depletion (Fig. 4b, c and Extended Data Fig. 8b, c), resulting in enhanced blood flow and significantly improved survival of ABX-treated SCD mice (Fig. 4d). Interestingly, the splenomegaly of SCD mice was significantly reduced, and liver damage including fibrosis, necrosis and inflammation was dramatically alleviated in ABX-treated SCD mice (Fig. 4e, f and Extended Data Fig. 8d, e). To test the impact of impaired neutrophil clearance in SCD, we depleted macrophages in the liver, spleen and bone marrow using clodronate liposomes¹⁷. Macrophage depletion markedly increased circulating aged neutrophils (data not shown) and resulted in acute vaso-occlusive crises that led to the death of all mice within 10–30 h (Fig. 4g). Together, these data suggest that the microbiota regulates aged neutrophil numbers, thereby affecting both acute vaso-occlusive crisis and the ensuing chronic tissue damage in SCD.

Finally, we evaluated whether the numbers of circulating aged neutrophils were altered in patients with SCD. As penicillin V antibiotic prophylaxis therapy is recommended for children less than 5 years old

or older patients with immune defects to prevent life-threatening infections³⁰, we determined aged neutrophil numbers in this patient population (Extended Data Fig. 8f, g). While there was no significant difference in total neutrophil numbers between SCD patients and healthy subjects, patients in the penicillin V prophylaxis group had significantly lower total neutrophil numbers (Fig. 4h). Consistent with our results in the mouse model, we found that SCD patients exhibited a marked increase in the numbers of circulating aged neutrophils compared to healthy controls. Remarkably, we observed significant reductions in both percentages and numbers of aged neutrophils in patients on penicillin V prophylaxis compared with SCD patients not taking antibiotics (Fig. 4h and Extended Data Fig. 8f). Age differences, gender or hydroxyurea intake in this case-control study did not mitigate the effect of antibiotic treatment on aged neutrophil numbers (Extended Data Fig. 8h), although a prospective study with age-matched subjects will be needed to ascertain the independent value of antibiotics in controlling aged neutrophil numbers in SCD.

Neutrophils are among the shortest-lived cells in the body¹. However, the evolutionary forces behind their rapid turnover remain unclear. Our results suggest that signals from the microbiota, through TLRs and Myd88, gradually lead them to become more functionally active. These data thus emphasize the notion that immunity is maintained by a balanced activation resulting from encounters with the microbiota, and provide a possible explanation for the evolutionary pressure that maintains an energy-consuming short lifespan as a mechanism to fine-tune the proportion of highly active neutrophils while balancing the risk of tissue injury. To our knowledge, this is the first therapy shown to alleviate the chronic tissue damage induced by SCD. Although antibiotic therapy normalized the overly active aged neutrophil population in SCD patients, the extent to which this treatment affects the gut microbiota or vaso-occlusive disease is open to future investigations. Our results raise the possibility that manipulation of the microbiome may have sustained implications in disease outcome that should be further studied in clinical trials.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 6 May; accepted 29 July 2015.

Published online 16 September; corrected online 23 September 2015

(see full-text HTML version for details).

1. Nauseef, W. M. & Borregaard, N. Neutrophils at work. *Nature Immunol.* **15**, 602–611 (2014).
2. Mayadas, T. N., Cullere, X. & Lowell, C. A. The multifaceted functions of neutrophils. *Annu. Rev. Pathol.* **9**, 181–218 (2014).
3. Taylor, P. R. *et al.* Activation of neutrophils by autocrine IL-17A-IL-17RC interactions during fungal infection is regulated by IL-6, IL-23, RORgammat and dectin-2. *Nature Immunol.* **15**, 143–151 (2014).
4. Woodfin, A. *et al.* The junctional adhesion molecule JAM-C regulates polarized transendothelial migration of neutrophils *in vivo*. *Nature Immunol.* **12**, 761–769 (2011).
5. Casanova-Acebes, M. *et al.* Rhythmic modulation of the hematopoietic niche through neutrophil clearance. *Cell* **153**, 1025–1035 (2013).
6. Rankin, S. M. The bone marrow: a site of neutrophil clearance. *J. Leukoc. Biol.* **88**, 241–251 (2010).
7. Martin, C. *et al.* Chemokines acting via CXCR2 and CXCR4 control the release of neutrophils from the bone marrow and their return following senescence. *Immunity* **19**, 583–593 (2003).
8. Stark, M. A. *et al.* Phagocytosis of apoptotic neutrophils regulates granulopoiesis via IL-23 and IL-17. *Immunity* **22**, 285–294 (2005).
9. Tanji-Matsuba, K. *et al.* Functional changes in aging polymorphonuclear leukocytes. *Circulation* **97**, 91–98 (1998).

10. Whyte, M. K., Meagher, L. C., MacDermot, J. & Haslett, C. Impairment of function in aging neutrophils is associated with apoptosis. *J. Immunol.* **150**, 5124–5134 (1993).
11. Hidalgo, A. *et al.* Heterotypic interactions enabled by polarized neutrophil microdomains mediate thromboinflammatory injury. *Nature Med.* **15**, 384–391 (2009).
12. Manwani, D. & Frenette, P. S. Vaso-occlusion in sickle cell disease: pathophysiology and novel targeted therapies. *Blood* **122**, 3892–3898 (2013).
13. Chiang, E. Y., Hidalgo, A., Chang, J. & Frenette, P. S. Imaging receptor microdomains on leukocyte subsets in live mice. *Nature Methods* **4**, 219–222 (2007).
14. Turhan, A., Weiss, L. A., Mohandas, N., Collier, B. S. & Frenette, P. S. Primary role for adherent leukocytes in sickle cell vascular occlusion: a new paradigm. *Proc. Natl Acad. Sci. USA* **99**, 3047–3051 (2002).
15. Scheiermann, C. *et al.* Adrenergic nerves govern circadian leukocyte recruitment to tissues. *Immunity* **37**, 290–301 (2012).
16. Gordy, C., Pua, H., Sempowski, G. D. & He, Y. W. Regulation of steady-state neutrophil homeostasis by macrophages. *Blood* **117**, 618–629 (2011).
17. Chow, A. *et al.* Bone marrow CD169⁺ macrophages promote the retention of hematopoietic stem and progenitor cells in the mesenchymal stem cell niche. *J. Exp. Med.* **208**, 261–271 (2011).
18. Subramanian, A. *et al.* Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl Acad. Sci. USA* **102**, 15545–15550 (2005).
19. Mortha, A. *et al.* Microbiota-dependent crosstalk between macrophages and ILC3 promotes intestinal homeostasis. *Science* **343**, 1249288 (2014).
20. Brenchley, J. M. & Douek, D. C. Microbial translocation across the GI tract. *Annu. Rev. Immunol.* **30**, 149–173 (2012).
21. Khosravi, A. *et al.* Gut microbiota promote hematopoiesis to control bacterial infection. *Cell Host Microbe* **15**, 374–381 (2014).
22. Deshmukh, H. S. *et al.* The microbiota regulates neutrophil homeostasis and host resistance to *Escherichia coli* K1 sepsis in neonatal mice. *Nature Med.* **20**, 524–530 (2014).
23. Balmer, M. L. *et al.* Microbiota-derived compounds drive steady-state granulopoiesis via MyD88/TICAM signaling. *J. Immunol.* **193**, 5273–5283 (2014).
24. Clarke, T. B. *et al.* Recognition of peptidoglycan from the microbiota by Nod1 enhances systemic innate immunity. *Nature Med.* **16**, 228–231 (2010).
25. Furze, R. C. & Rankin, S. M. The role of the bone marrow in neutrophil clearance under homeostatic conditions in the mouse. *FASEB J.* **22**, 3111–3119 (2008).
26. Sabroe, I. *et al.* Selective roles for Toll-like receptor (TLR)2 and TLR4 in the regulation of neutrophil activation and life span. *J. Immunol.* **170**, 5268–5275 (2003).
27. Magri, G. *et al.* Innate lymphoid cells integrate stromal and immunological signals to enhance antibody production by splenic marginal zone B cells. *Nature Immunol.* **15**, 354–364 (2014).
28. Li, P. *et al.* PAD4 is essential for antibacterial innate immunity mediated by neutrophil extracellular traps. *J. Exp. Med.* **207**, 1853–1862 (2010).
29. Chen, G. *et al.* Heme-induced neutrophil extracellular traps contribute to the pathogenesis of sickle cell disease. *Blood* **123**, 3818–3827 (2014).
30. Gaston, M. H. *et al.* Prophylaxis with oral penicillin in children with sickle cell anemia. A randomized trial. *N. Engl. J. Med.* **314**, 1593–1599 (1986).

Acknowledgements We are grateful to C. Prophete and P. Ciero for expert technical assistance. We also thank E. Pamer (Memorial Sloan Kettering Cancer Center) for the gift of *Tlr2*^{-/-} and *Tlr4*^{-/-} mice; K. Ireland for assistance with the SCD patient study; Z. Chen for the taxonomic microbiota analysis; R. Ng for assistance with the germ-free mice; O. Uche and G. Wang for assistance in cell sorting; D. Reynolds and W. Tran for the microarray assay; and R. Sellers for histopathological analyses. This work was supported by a predoctoral fellowship from the American Heart Association (15PRE23010014 to D.Z.) and R01 grants from the National Institutes of Health (HL069438, DK056638, HL116340 to P.S.F.). Flow cytometry and cell sorting was supported by a Shared Facilities Award from the New York State Stem Cell Science (NYSTEM) Program.

Author Contributions D.Z. designed and performed experiments, analysed results and wrote the manuscript; G.C., C.X., Y.K., R.B. and J.J. performed experiments and provided valuable inputs on the manuscript; A.M. provided *LysM-cre/Myd88*^{fl/fl} and *Csf2*^{-/-} mice and performed experiments; J.J.F. provided germ-free mice and performed experiments; D.M. provided human samples; C.S. and M.M. discussed data and provided valuable input on the manuscript; P.S.F. designed and supervised the study, discussed data and wrote the manuscript.

Author Information Microarray data have been deposited in the Gene Expression Omnibus under accession code GSE69886. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to P.S.F. (paul.frenette@einstein.yu.edu).

METHODS

Mice. *Selp*^{-/-}, *CD169-DTR*, *Csf2*^{-/-} mice, Tg[*Hu-miniLCRα1*^{G_γA_γδβ^S}] *Hba*^{-/-} *Hbb*^{-/-} (Berkeley sickle-cell mice) and Tg[*Hu-miniLCRα1*^{G_γA_γδβ^S}] *Hba*^{-/-} *Hbb*^{+/-} (hemizygous control mice) have been described previously^{11,14,17,19}. B6.129P2-*Lyz2*^{tm1(cres)lfo}/J (*LysM-cre*), B6.129P2(SJL)-*Myd88*^{tm1Defr}/J (*Myd88*^{fl/fl}) and B6.129S-*Tnf*^{tm1Gkl}/J (*Tnf*^{-/-}) mice were purchased from The Jackson Laboratory. *Tlr2*^{-/-} and *Tlr4*^{-/-} mice were kindly provided by E. G. Pamer (Memorial Sloan-Kettering Cancer Center, NY). C57BL/6 CD45.1 and CD45.2 mice were purchased from the National Cancer Institute. Six- to eight-week-old mice were used for experiments. All mice were housed in specific-pathogen-free conditions and fed with autoclaved food; experimental procedures performed on mice were approved by the Animal Care and Use Committee of Albert Einstein College of Medicine. Germ-free C57BL/6 mice were maintained in sterile isolators with autoclaved food and water in the Gnotobiotic Core of Icahn School of Medicine at Mount Sinai. For fecal transplantation experiments, 100 mg of feces pellets was resuspended in 1 ml of PBS, homogenized, and filtered through a 70-μm strainer. Recipient germ-free mice were gavaged with 200 μl of the filtrate.

Human samples. Blood was obtained from healthy volunteers, SCD patients and SCD patients on penicillin V prophylaxis after parental consent and child assent as approved by the Institutional Review Board of Albert Einstein College of Medicine. SCD patients were recruited upon routine visits at the sickle-cell clinic of Montefiore Medical Center. Among the 34 patients recruited for the study, 11 were on penicillin V owing to age (less than 5 years old) or defective immunity, and 23 were off antibiotic treatment for at least 2 months. Patients with acute infection or vaso-occlusive crisis were excluded from the study.

Bone marrow transplantation. Age- and gender-matched sickle-cell disease (SS) and control hemizygous (SA) mouse cohorts were generated by transplanting bone marrow nucleated cells from Berkeley sickle-cell mice or control hemizygous mice into lethally irradiated C57BL/6 mice as described before^{11,14}. Fully reconstituted mice (>97%) were used for studies. Chimaeric mice used to study neutrophil TLRs were generated by transplanting a 1:1 mixture of bone marrow nucleated cells from C57BL/6 mice (CD45.1⁺) and *LysM-cre/Myd88*^{fl/fl} or *Tlr4*^{-/-} or *Tlr2*^{-/-} mice (CD45.2⁺) into lethally irradiated C57BL/6 recipients (CD45.1⁺). Chimaeric mice were analysed 6 weeks after transplantation.

Antibiotic treatment. Wild-type or SCD mice were treated with ampicillin (1 g l⁻¹), neomycin (1 g l⁻¹), metronidazole (1 g l⁻¹) and vancomycin (1 g l⁻¹) in drinking water for 4–6 weeks. Antibiotics were purchased from Sigma or Jack D. Weiler Hospital of the Albert Einstein College of Medicine. Drinking water containing antibiotics was changed every 3–4 days. For microbial product reintroduction experiments, ABX-treated wild-type mice were fed with 1 mg LPS (0111:B4, Sigma), 1 mg peptidoglycan (PGN-SA, Invivogen) or 1 mg MurNAc-L-Ala-γ-D-Glu-mDAP (M-TriDAP, Invivogen) by intragastric gavage, and allowed to rest for 24–36 h. For the analysis of neutrophil–LPS interaction, untreated wild-type mice were fed with 300 μg LPS–FITC (Sigma) by intragastric gavage, and tissues were harvested 1 h after gavage.

Adoptive transfer. For *in vivo* neutrophil ageing analysis, whole blood from donor mice was transfused into recipient mice by retro-orbital injection. Donor neutrophils in blood were tracked based on CD45.1 and CD45.2 expression by flow cytometry. For microarray analyses, control and aged neutrophils were derived by *in vivo* ageing for 10 min and 6 h, respectively. Activated neutrophils were harvested from mice injected with 0.5 μg TNFα (R&D Systems) for 2 h. To analyse Mac-1 activation of neutrophils from wild-type and *Selp*^{-/-} mice, 3–5 × 10⁶ leukocytes were harvested from blood following red blood cell (RBC) lysis, and were labelled with red fluorescent dye PKH26 (Sigma) according to the manufacturer's protocol before the transfer into recipient mice.

Macrophage depletion. For the depletion of CD169⁺ macrophages, wild-type or *CD169-DTR* mice were injected intraperitoneally with two doses of 10 μg kg⁻¹ body weight diphtheria toxin (Sigma) 3 days apart. Mice were analysed 5 days after the second injection. For depletion of macrophages in SCD mice, mice were injected intravenously with 250 μl PBS- or clodronate-encapsulated liposomes (the Foundation Clodronate Liposomes) as described before¹⁷.

Flow cytometry and cell sorting. Cells were surface-stained in PEB buffer (PBS supplemented with 0.5% BSA and 2 mM EDTA) for 20–30 min on ice. Multiparametric flow cytometric analyses were performed on a LSRII equipped with FACS Diva 6.1 software (BD Biosciences) and analysed with FlowJo software (Tree Star). Dead cells were excluded by FSC, SSC and 4',6-diamino-2-phenylindole (DAPI, Sigma) staining. Cell sorting experiments were performed on Aria Cell Sorter (BD Biosciences). Neutrophils were gated by Gr-1^{hi} CD115^{lo} SSC^{hi}; T cells, B cells and monocytes were gated by CD3⁺, CD4⁺ and CD115^{hi}; aged neutrophils were gated by CD62L^{lo} CXCR4^{hi} within the neutrophil population; haematopoietic progenitor and stem cells were identified by lineage cocktail, Sca-1, KitL, CD150, CD48, CD34 and CD16/32, as previously described³¹; macrophages were identified by Gr-1^{lo} CD115^{lo} F4/80⁺ SSC^{lo} as previously described¹⁷.

Fluorophore-conjugated or biotinylated antibodies against mouse Gr-1 (RB6-8C5), CD115 (AFS98), CD3 (145-2C11), B220 (RA3-6B2), PE-anti-CXCR4 (2B11), CD45.1 (A20), CD45.2 (104), CD11b (M1/70), ICAM-1 (YN1/1.7.4), CD11c (N418), CD49d (R1-2), CD45 (30-F11), Sca-1 (D7), c-Kit (2B8), CD34 (RAM34), and F4/80 (BM8) were from eBioscience. Antibodies specific to CD62L (MEL-14) and Biotin Mouse Lineage Panel (TER-119, RB6-8C5, RA3-6B2, M1/70, 145-2C11) were from BD Pharmingen. Antibodies against CD47 (miap301), CD150 (TC15-12F12.2), CD48 (HM48-1) and CD16/32 (93) were from BioLegend.

Blood leukocyte counts. Blood was harvested from the retro-orbital plexus, collected in EDTA or heparin, and analysed on an ADVIA 120 hematology system (SIEMENS).

Brightfield intravital microscopy. Experimental procedures and data analyses were performed as previously described^{11,13,14}. In brief, male mice were injected intrascrotally with 0.5 μg TNFα (R&D Systems), and were anaesthetized 2 h later by intraperitoneal injection of a mixture of 2% chloralose (Sigma) and 10% urethane (Sigma) in PBS. Tracheal intubation was performed to ensure normal respiration after anaesthesia. The cremaster muscle was gently exteriorized, and mounted onto a microscopic stage, and then superfused with Ringer solution (pH 7.4, 37 °C). Under the microscope, leukocyte rolling, adhesion, transmigration and neutrophil–RBC interactions in post-capillary venules (20–40 μm in diameter) were captured using a custom-designed upright microscope (MM-40, Nikon) equipped with a 60 × water-immersion objective (Nikon). Adhesion was quantified as the number of leukocytes remaining stationary for more than 20 s within a 100 μm venular segment. In this model, more than 90% of adherent leukocytes are neutrophils based on previous studies¹³. Neutrophil–RBC interactions were defined as the associations between an RBC and an adherent leukocyte for more than 3 s, and quantified as the number of interactions within a 100 μm vessel segment per minute. At least eight vessels from each animal were recorded and analysed using a charge-coupled video camera (Hamamatsu) and video recorder (Sony SVHS, SVO-9500). Venular diameters were measured using a video caliper, and centerline red cell velocity (*V*_{RBC}) for each venule recorded was measured using an optical Doppler velocimeter (Texas A&M). Blood flow rate (*Q*) was calculated as $Q = \frac{V_{\text{mean}} \times \pi d^2}{4}$, where *d* is venule diameter, *V*_{mean} is estimated as $\frac{V_{\text{RBC}}}{1.6}$.

Multi-channel fluorescence intravital microscopy. MFIM analyses of Mac-1 activation of adherent neutrophils were performed as previously described¹¹. In brief, yellow-green fluorescent microsphere beads (1.0 μm, Life Technologies) were incubated with 1 mg ml⁻¹ BSA (Fisher Bioreagents) for at least 2 h in PBS and sonicated for 15 min in a water-bath sonicator (Laboratory Supplies Co.). Albumin-coated beads (10⁹ beads per mouse) were injected into mice prepared for intravital microscopy 3 h after TNFα injection. For measurement of CD62L expression on adherent neutrophils, mice were intravenously injected with 4 μg APC-anti-CD62L (clone MEL-14, BD Pharmingen). For *in vivo* staining of CD45 alleles, mice were intravenously injected with 5–10 μg Alexa Fluor 647-anti-CD45.2 (clone 104, Biolegend) and biotin-anti-CD45.1 (clone A20, eBioscience), and 20 min later injected with 5–10 μg streptavidin eFluor 450 (eBioscience). Images and videos were captured using an Axio Examiner.D1 microscope (Zeiss) equipped with a Yokogawa CSU-X1 confocal scan head with four stack laser system (405 nm, 488 nm, 561 nm, and 642 nm wavelengths) and a 60 × water-immersion objective, and analysed using Slidebook software (Intelligent Imaging Innovations). Mac-1 activation of adherent neutrophils was quantified as the average number of beads captured by each adherent neutrophil in post-capillary venules, and the percentage of cells that captured more than eight beads.

Intravital microscopic studies of SCD mice. Male mice were injected intraperitoneally with 0.5 μg TNFα (R&D Systems), and 2 h later neutrophil responses were analysed as described above. Survival times, defined as the time from TNFα injection until death, were recorded.

Microarray analysis. Total RNA from 2,000 sorted neutrophils was extracted using RNeasy Plus Micro Kit (Qiagen) according to the manufacturer's protocol. All further steps were performed at the Genomics Core Facility at Albert Einstein College of Medicine. RNA was amplified using Ovation One-Direct System (NuGEN). Amplified cRNA was labelled with the GeneChip wild-type terminal labelling kit (Affymetrix), hybridized to Mouse Gene ST 1.0 microarrays (Affymetrix), and scanned by GeneChip Scanner 3000 7G system (Affymetrix) according to standard protocols. Raw data were normalized by RMA algorithm and analysed using the Gene Pattern analysis platform (Broad Institute). After removal of unannotated genes, gene set enrichment analysis was performed with all C2 gene sets from the Molecular Signatures Database (v5.0, Broad Institute). Gene sets with a *P* value of <0.05 in either aged or activated groups were considered to have significant differences compared to control group. Normalized enrichment score (NES) for selected pathways related to neutrophil functions were depicted as a heat map, with gene sets clustered by functional classifications.

Quantitative real-time PCR (qPCR). Messenger RNA extraction from 500–2,000 sorted neutrophils using Dynabeads mRNA DIRECT Kit (Life Technologies) and reverse transcription using RNA to cDNA EcoDry Premix (Clontech) were performed according to the manufacturer's protocols. qPCR was performed with SYBR GREEN (Roche) on ABI PRISM 7900HT Sequence Detection System (Life Technologies). The PCR protocol started with one cycle at 95 °C (10 min) and continued with 40 cycles at 95 °C (15 s) and 60 °C (1 min). Expression of glyceraldehyde-3-phosphate dehydrogenase (*Gapdh*) was used as a standard. The average threshold cycle number (C_T) for each tested gene was used to quantify the relative expression of each gene: $2^{(C_{T(\text{standard})} - C_{T(\text{gene})})}$. Primers include: *Gapdh* forward, TGTGTCCGTCGTGGATCTGA; *Gapdh* reverse, CCTGC TTCACCACTTCTTGA; *Tlr4* forward, ATGGCATGGCTTACACCAC; *Tlr4* reverse, GAGGCCAATTTGTCTCCACA; *Icam1* forward, GGACCAC GGAGCCAATTTTC; *Icam1* reverse, CTCGGAGACATTAGAGAACAATGC; *Itgam* forward, CTGAACATCCCATGACCTTCC; *Itgam* reverse, GCCCAAGG ACATATTACAGC; *Cxcl2* forward, CGCTGTCAATGCCTGAAG; *Cxcl2* reverse, GCGGTCACACTCAAGCTCT.

ELISA. Concentrations of IFN- γ , IL-6, TNF α and IL-1 β were measured in plasma from wild-type and *CD169-DTR* mice 5 days after diphtheria toxin treatment using ELISA kits (eBioscience) according to the manufacturer's instructions. For measurement of NET biomarkers, plasma nucleosome was measured using Cell Death Detection ELISA Plus kit (Roche), and plasma DNA was measured using Sytox Green (Life Technologies) as previously described²⁹.

16S rDNA quantification. Stool pellets were collected and total DNA was extracted using the QIAamp Fast DNA Stool Mini Kit (Qiagen). Quantification of 16S rDNA was performed by real-time qPCR using TaqMan Universal Master Mix (Life Technologies) and the following primers and probe as described³²; forward, 5'-ACTGAGAYACGGYCCA-3'; reverse, 5'-TTACCGCGGCTGC TGGC-3'; Probe 6-FAM-ACTCCTACGGGAGGCAGCAGT-BHQ1.

Taxonomic microbiota analysis. Taxonomic microbiota analysis was performed by the Molecular Biology and Next Generation Technology Core at Albert Einstein College of Medicine. In brief, purified 16S rDNA was used for PCR amplification and sequencing. The variable region 4–6 (V4–V6) of the 16S rDNA gene was amplified using barcoded 16V6R1052 and 16SV4F515 primers. Sequencing was performed using paired-end Illumina MiSeq sequencing. Taxonomical classification was obtained using the RDP-classifier to generate an OTU table, and the percentage of each family genus in total microbiome was derived from the OTU values.

Neutrophil release–clearance kinetics. Mice were injected intraperitoneally with 2 mg EdU and were bled on day 2–7 after EdU injection. Each mouse was bled only once to avoid the potential change in kinetics caused by bleeding. Cells were surface stained, fixed with 2% paraformaldehyde (PFA), and permeabilized with 0.1% Triton-X. After permeabilization, EdU incorporation was detected by Click-iT EdU Alexa Fluor 647 Imaging Kit (Life Technologies) according to the manufacturer's instructions.

In vitro NET assay. Circulating neutrophils were harvested using Percoll Density Centrifugation Media (GE Healthcare) as previously described²⁹. In brief, blood was loaded onto a Percoll gradient consisting of 52%, 65% and 78% Percoll layers, and centrifuged at 2,500 r.p.m. for 30 min at room temperature. The cell bands between 65% and 78% layers were harvested, and RBCs were removed using RBC lysis buffer. Purity of 80–95% was constantly achieved with this method, as analysed by flow cytometry. For ROS production assay, neutrophils were treated with 20 $\mu\text{g ml}^{-1}$ LPS (0111:B4, Sigma) for 30 min. ROS detection was performed using fluorogenic dye Dihydrorhodamine 123 (Life Technologies) as previously described³³. For NET formation *in vitro*, neutrophils were attached to poly-L-lysine-coated slides and treated with 20 $\mu\text{g ml}^{-1}$ LPS (0111:B4, Sigma) for 3 h. Following stimulation, cells were stained without fixation with SYTOX Orange (cell impermeable) and SYTO 13 (cell permeable) nucleic acid dyes (Life Technologies) to image DNA fibres and distinguish live and dead cells. After DNA staining, cells were fixed, permeabilized and blocked as previously described²⁹. Cells were incubated with goat anti-neutrophil elastase (sc-9521, Santa Cruz Biotechnology) and rabbit anti-CitH3 (ab5103, Abcam) followed by Alexa Fluor 647 Donkey-anti-goat (Life Technologies) and Brilliant Violet 421 donkey anti-rabbit (Biolegend) secondary antibodies. Species-matching isotype controls were used to confirm fluorescence staining. NETs were defined by DNA

fibres co-localized with neutrophil elastase and CitH3 proteins with a length exceeding 40 μm , and quantified as the percentage of NETs among all neutrophils present in the field.

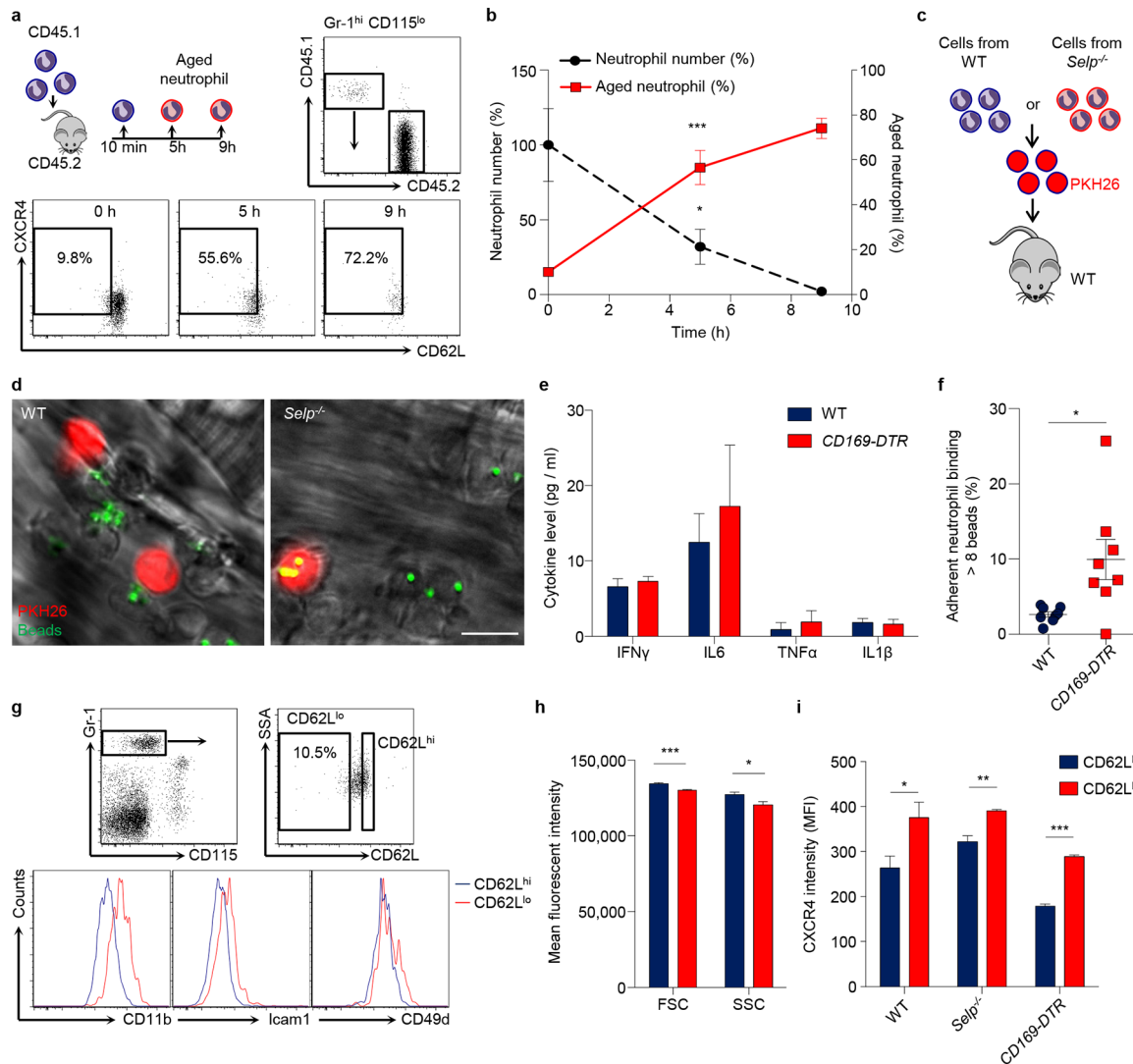
In vivo NET assay. For analyses of NET formation *in vivo*, goat anti-neutrophil-elastase (sc-9521, Santa Cruz Biotechnology), and rabbit anti-CitH3 (ab5103, Abcam) were labelled by APEX Alexa Fluor 647 and APEX Alexa Fluor 568 antibody labelling kit (Life Technologies), respectively, according to the manufacturer's instructions. Mice were injected intraperitoneally with 30 mg kg^{-1} LPS for 3 h, and then were injected intravenously with 5 μg Alexa Fluor 568-labelled anti-CitH3, 2 μg Alexa Fluor 647-labelled anti-NE, 10 μg Pacific Blue-labelled anti-CD31 (clone 390 or clone MEC13.3, Biolegend) and 10 μM SYTOX Green (cell impermeable) nucleic acid dye (Life Technologies). Species-matching isotype controls were also labelled using the same protocol and injected into septic mice to validate fluorescence staining. Fresh livers were obtained 20 min after antibody injections, and directly imaged using an Axio Examiner.D1 confocal microscope (Zeiss). Confocal microscopy provided a penetration of $\sim 100 \mu\text{m}$ into the tissue. Liver vasculature was identified by CD31 staining. NETs were defined by extracellular DNA fibres stained by SYTOX Green, anti-CitH3 and anti-neutrophil-elastase antibodies with a length exceeding 40 μm , and quantified as the average number of NETs in each vessel.

Immunofluorescence. Tissues were embedded in Tissue-Plus O.C.T. Compound (Fisher HealthCare), and frozen 20- μm thick sections were prepared using a CM3050 S cryostat (Leica). Sections were fixed with 4% PFA for 10 min, and blocked and permeabilized with PBS containing 20% species-matching serum and 0.5% Triton-X for 1–2 h. Adhesion molecules on endothelial cells were stained by PE-anti-P-selectin (clone Psel.KO2.3, eBioscience), PE-anti-E-selectin (clone 10E9.6, BD Pharmingen), PE-anti-Icam1 (clone YN1/1.7.4, eBioscience), and PE-anti-VCAM-1 (clone 429, Biolegend). Expression of adhesion molecules was quantified using Slidebook software (Intelligent Imaging Innovations) as previously described¹⁵. For immunofluorescence staining of neutrophils on sections, goat anti-neutrophil-elastase (sc-9521, Santa Cruz Biotechnology) and rabbit anti-CitH3 (ab5103, Abcam) were used, followed by Alexa Fluor 568 Donkey-anti-goat and Alexa Fluor 488 donkey anti-rabbit (Life Technologies) secondary antibodies. For analyses of fibrin deposition, sections were fixed using formalin containing 2% acetic acid for 30 min to remove soluble fibrinogen and leave only cross-linked fibrin in the tissue³⁴. Sections were then blocked and permeabilized, and incubated with goat anti-fibrinogen- β (sc-18029, Santa Cruz Biotechnology) and PE-anti-Ly6G (clone 1A8, biolegend), followed by Alexa Fluor 568 donkey-anti-goat secondary antibody (Life Technologies). In all immunofluorescence experiments, endothelial cells were identified by Alexa Fluor 647-anti-CD31 (MEC13.3, Biolegend) and nuclei were stained by Hoechst 33342 (Life Technologies).

Histology analyses of septic mice. Mice were injected intraperitoneally with 30 mg kg^{-1} LPS, and livers were collected 24 h after injection and fixed in 10% formalin. Histological analyses were performed by the Histology and Comparative Pathology Facility at Albert Einstein College of Medicine according to standard protocols. Survival time was defined as the time from LPS injection until death of the mouse to a maximum of 150 h.

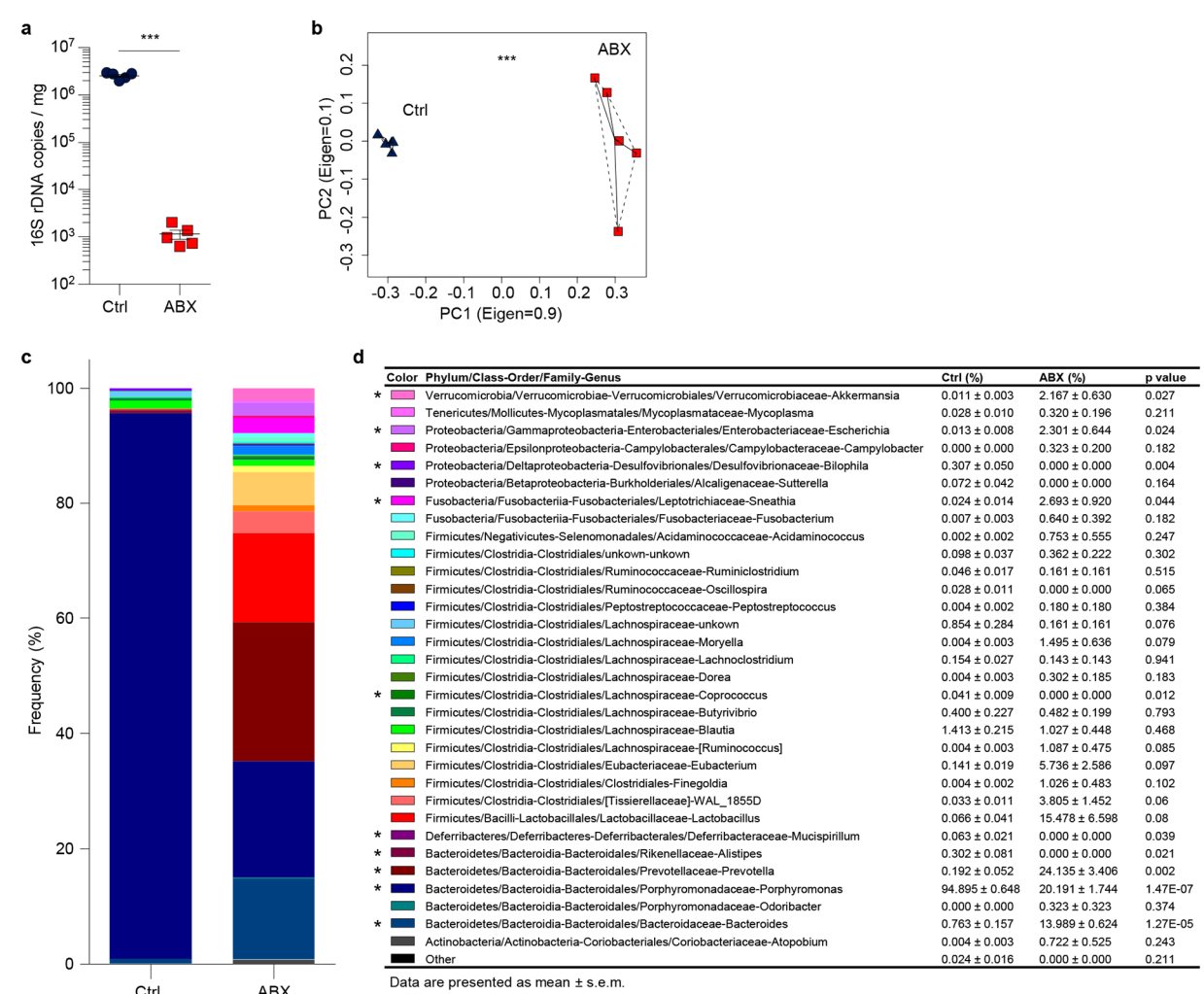
Statistical analysis. No statistical methods were used to predetermine sample size. Experiments were performed blind to group allocations, and validated by two investigators independently. Paired and unpaired two-tailed Student's *t*-tests and Mann–Whitney *U*-tests were used to compare two groups. One-way ANOVA analysis was used for multiple group comparisons. Log-rank test was used to compare survival curves. Statistical analyses were performed using Graph Pad Prism 6 software.

- Hanoun, M. *et al.* Acute myelogenous leukemia-induced sympathetic neuropathy promotes malignancy in an altered hematopoietic stem cell niche. *Cell Stem Cell* **15**, 365–375 (2014).
- Hang, J. *et al.* 16S rRNA gene pyrosequencing of reference and clinical samples and investigation of the temperature stability of microbiome profiles. *Microbiome* **2**, 31 (2014).
- Rothe, G. & Valet, G. Flow cytometric assays of oxidative burst activity in phagocytes. *Methods Enzymol.* **233**, 539–548 (1994).
- Schnitt, S. J. *et al.* Myocardial fibrin deposition in experimental viral myocarditis that progresses to dilated cardiomyopathy. *Circ. Res.* **72**, 914–920 (1993).



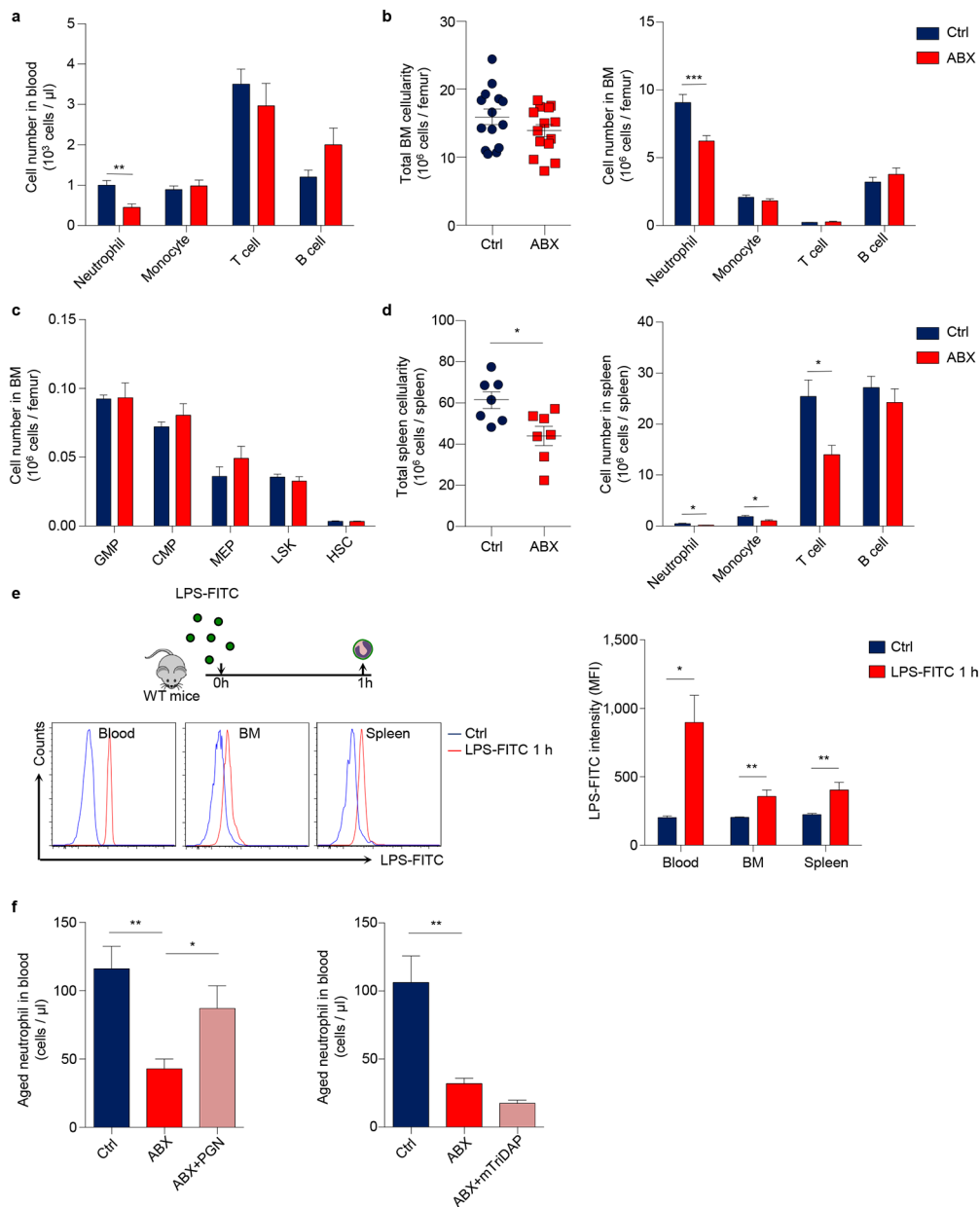
Extended Data Figure 1 | Phenotypic and functional characterization of aged neutrophils. **a**, Flow cytometry analysis of donor neutrophil ageing after adoptive transfer into recipients. Donor neutrophils gated by CD45.1⁺ and aged neutrophils gated by CD62L^{lo}CXCR4^{hi}. **b**, Ageing and clearance kinetics of donor neutrophils after adoptive transfer into recipients ($n = 3$ mice). Left y axis, donor neutrophil number relative to the initial number of neutrophils transferred (black dashed line); right y axis, percentage of the aged subset in donor neutrophils (red line). **c**, **d**, MFIM analysis of Mac-1 activation of neutrophils harvested from wild-type or *Selp*^{-/-} mice, labelled by PKH26 (red) and transferred into wild-type recipients. Scale bar, 10 μ m. **e**, Plasma cytokine levels in wild-type and *CD169-DTR* mice 5 days after diphtheria toxin

treatment ($n = 5$ mice). **f**, Percentages of adherent neutrophils that capture more than eight beads in diphtheria-toxin-treated wild-type and *CD169-DTR* mice ($n = 8$ mice). **g**, **h**, Flow cytometry analysis of surface marker expression (**g**), cell size (FSC) and granularity (SSC; **h**; $n = 7$ mice) on CD62L^{hi} young and CD62L^{lo} aged neutrophils. **i**, CXCR4 expression levels on CD62L^{hi} young and CD62L^{lo} aged neutrophils in wild-type, *Selp*^{-/-}, and *CD169-DTR* mice (wild type, $n = 13$ mice; *Selp*^{-/-}, $n = 4$ mice; *CD169-DTR*, $n = 5$ mice). Error bars, mean \pm s.e.m. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$, data representing two or more independent experiments analysed with one-way ANOVA (**b**) or unpaired Student's *t*-test (**e-i**).



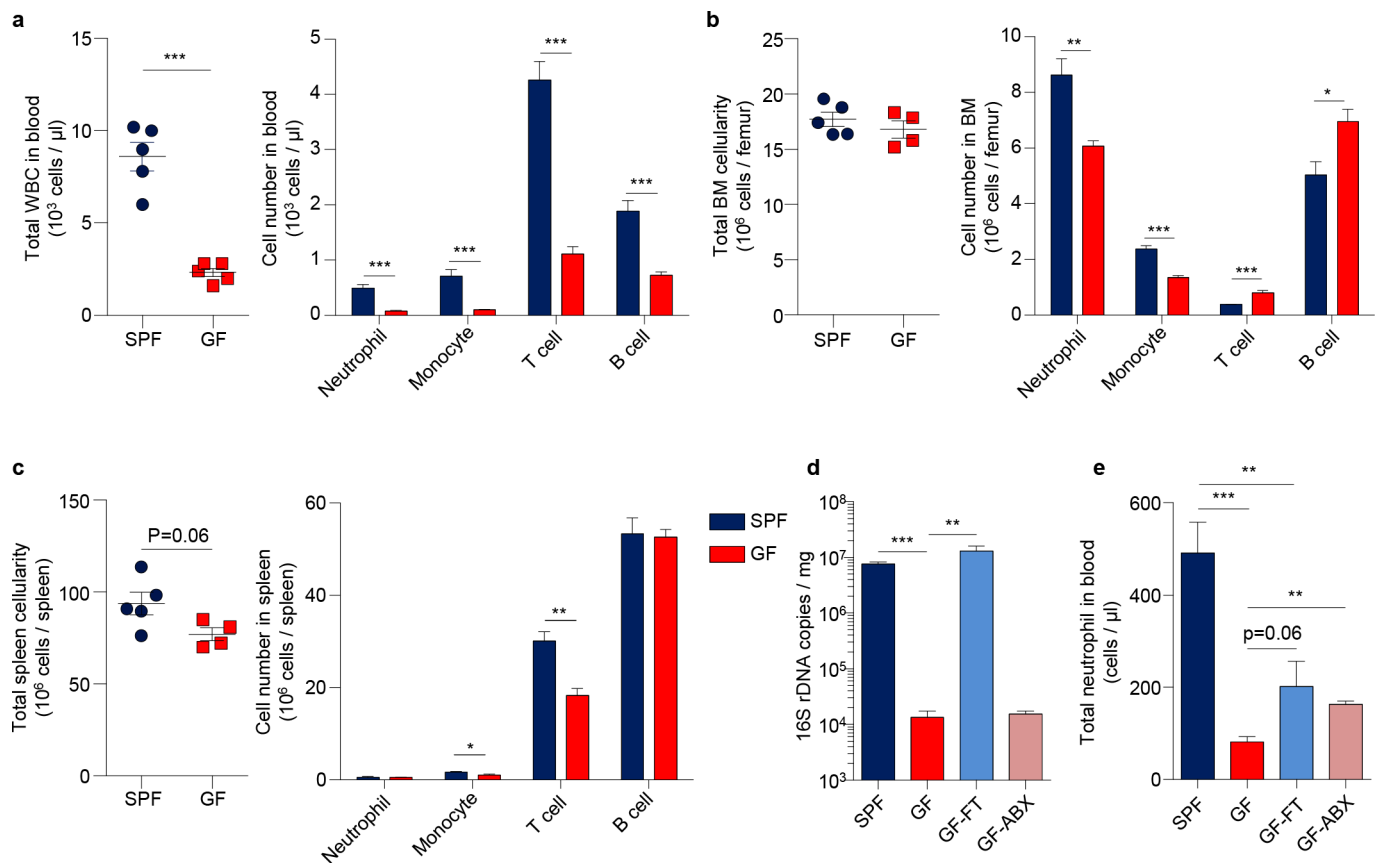
Extended Data Figure 2 | Antibiotic treatment efficiently depletes and alters the composition of the microbiota. **a**, Copy numbers of 16S ribosomal DNA in feces from control and antibiotics (ABX)-treated mice ($n = 5$ mice). **b**, Principal component analysis of the microbiome composition in control and ABX-treated mice ($n = 5$ mice). **c**, **d**, Percentage of each bacteria genus in

total microbiome ($n = 5$ mice). Error bars, mean ± s.e.m. $*P < 0.05$, $***P < 0.001$, data representing two or more independent experiments analysed with unpaired Student's t -test (**a**, **d**) or permutational multivariate ANOVA (**b**).



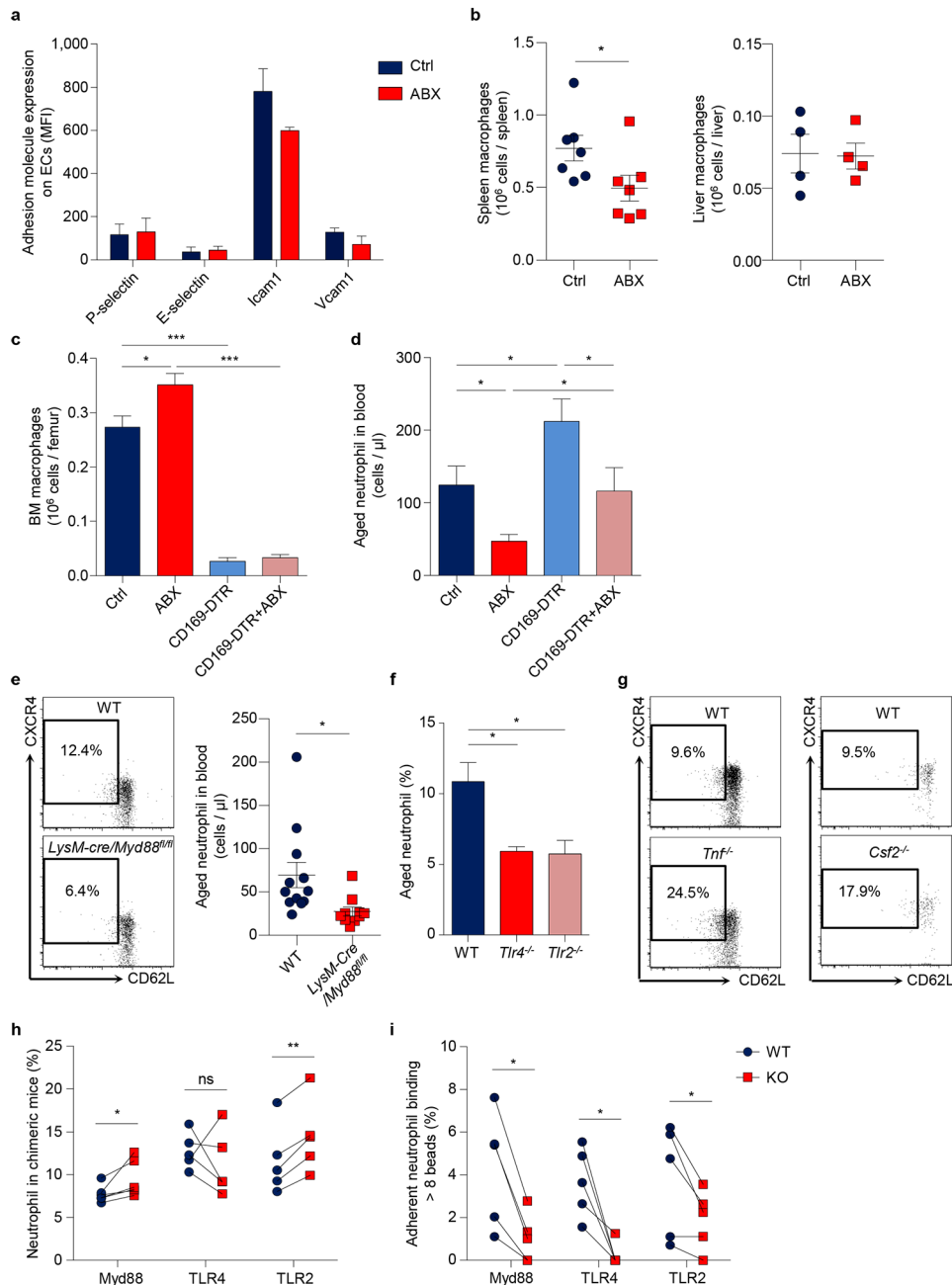
Extended Data Figure 3 | Microbiota-derived molecules regulate neutrophil homeostasis and ageing. **a**, Numbers of circulating leukocyte subsets in control and antibiotics (ABX)-treated mice ($n = 9$ mice). **b**, Bone marrow cellularity and numbers of leukocyte subsets in the bone marrow of control and ABX-treated mice ($n = 14$ mice). **c**, Numbers of bone marrow haematopoietic stem and progenitor cells in control and ABX-treated mice ($n = 9$ mice). **d**, Spleen cellularity and numbers of leukocyte subsets in the spleen of control and ABX-treated mice ($n = 7$ mice). **e**, Flow cytometry analysis of neutrophil-LPS interactions in blood, bone marrow (BM) and spleen 1 h after LPS-FITC

gavage (Ctrl, $n = 4$ mice; LPS-FITC, $n = 5$ mice). Histogram showing fluorescence intensity on neutrophils gated by Gr-1^{hi} CD115^{lo} SSA^{hi}. **f**, Numbers of circulating aged neutrophils in control, ABX-treated, and ABX-treated mice fed with peptidoglycan (PGN) or mTriDAP (left, $n = 11$ (Ctrl), 9 (ABX), 9 (ABX+PGN) mice; right, $n = 10$ (Ctrl), 10 (ABX), 5 (ABX+mTriDAP) mice). Error bars, mean \pm s.e.m. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$, data representing two or more independent experiments analysed with unpaired Student's t -test.



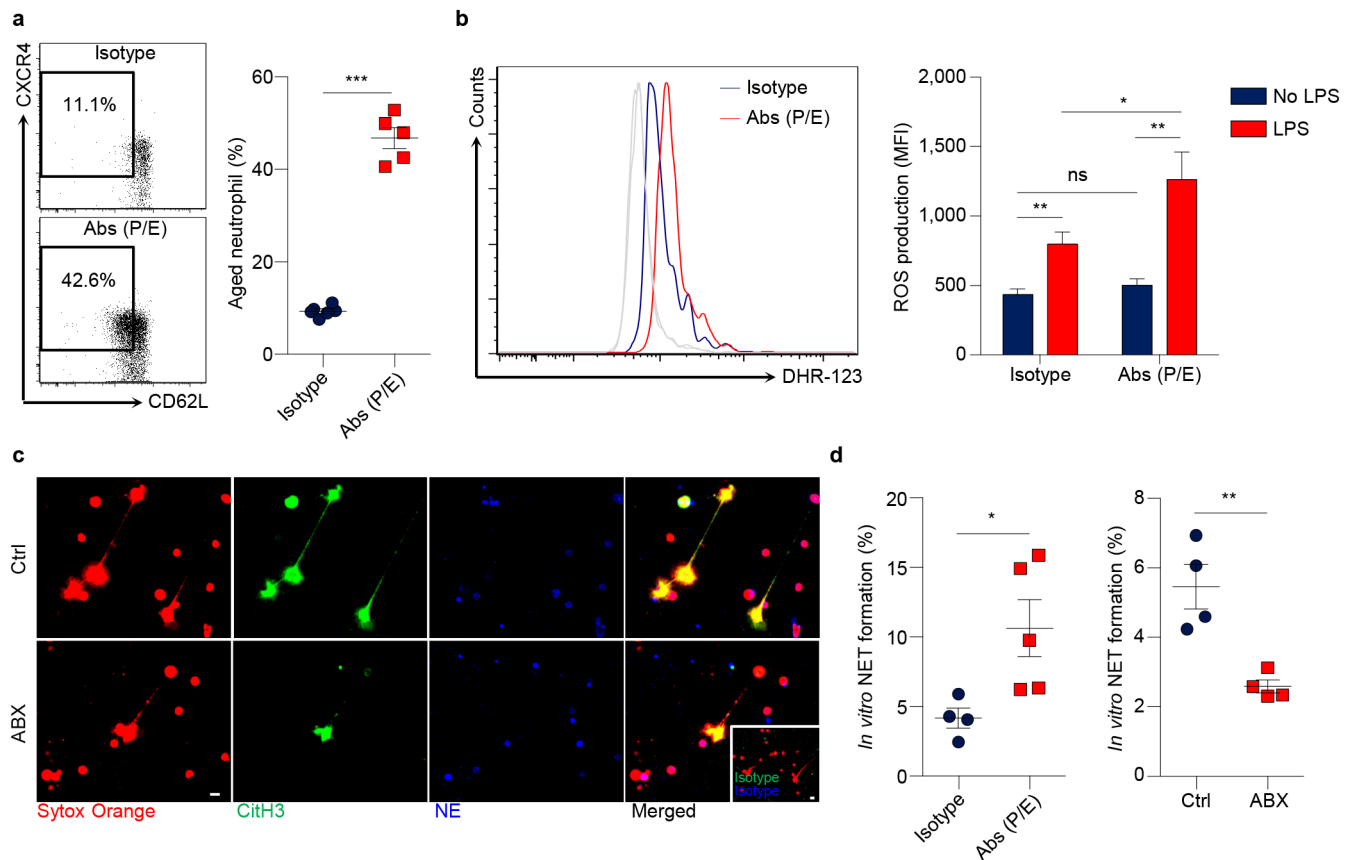
Extended Data Figure 4 | Neutrophil homeostasis is altered in germ-free mice. **a**, Total white blood cell (WBC) counts and numbers of leukocyte subsets in blood of specific-pathogen-free (SPF) and germ-free (GF) mice ($n = 5$ mice). **b**, Total bone marrow (BM) cellularity and numbers of leukocyte subsets in the bone marrow of SPF and germ-free mice (SPF, $n = 5$ mice; germ-free, $n = 4$ mice). **c**, Total spleen cellularity and numbers of leukocyte subsets in the spleen of SPF and germ-free mice (SPF, $n = 5$ mice; germ-free, $n = 4$ mice).

d, Copy numbers of 16S ribosomal DNA in feces from SPF mice, germ-free mice, germ-free mice reconstituted by fecal transplantation (GF-FT), and antibiotic-treated germ-free mice (GF-ABX; $n = 5, 5, 5$ and 4 mice, respectively). **e**, Numbers of total circulating neutrophils in SPF, germ-free, GF-FT, and GF-ABX mice ($n = 5, 5, 5$ and 3 mice, respectively). Error bars, mean \pm s.e.m. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$, data representing two or more independent experiments analysed with unpaired Student's t -tests.



Extended Data Figure 5 | Microbiota-driven neutrophil ageing is independent of clearance mechanisms, and mediated by TLRs and Myd88 signalling. **a**, Adhesion molecule expression on endothelial cells (ECs) in control and antibiotics (ABX)-treated mice ($n = 4$ mice). MFI, mean fluorescence intensity. **b**, Numbers of spleen and liver macrophages in control and ABX-treated mice (left, $n = 7$ mice; right, $n = 4$ mice). **c**, **d**, Numbers of bone marrow (BM) macrophages (**c**; $n = 19, 19, 10, 10$ mice, left to right) and circulating aged neutrophils (**d**; $n = 12, 11, 10, 9$ mice, left to right) in diphtheria toxin (DT)-treated control, ABX-treated mice, *CD169-DTR*, and ABX-treated *CD169-DTR* mice. **e**, Flow cytometry analysis of aged neutrophils in wild-type and *LysM-cre/Myd88^{fl/fl}* mice ($n = 12, 10$ mice, respectively).

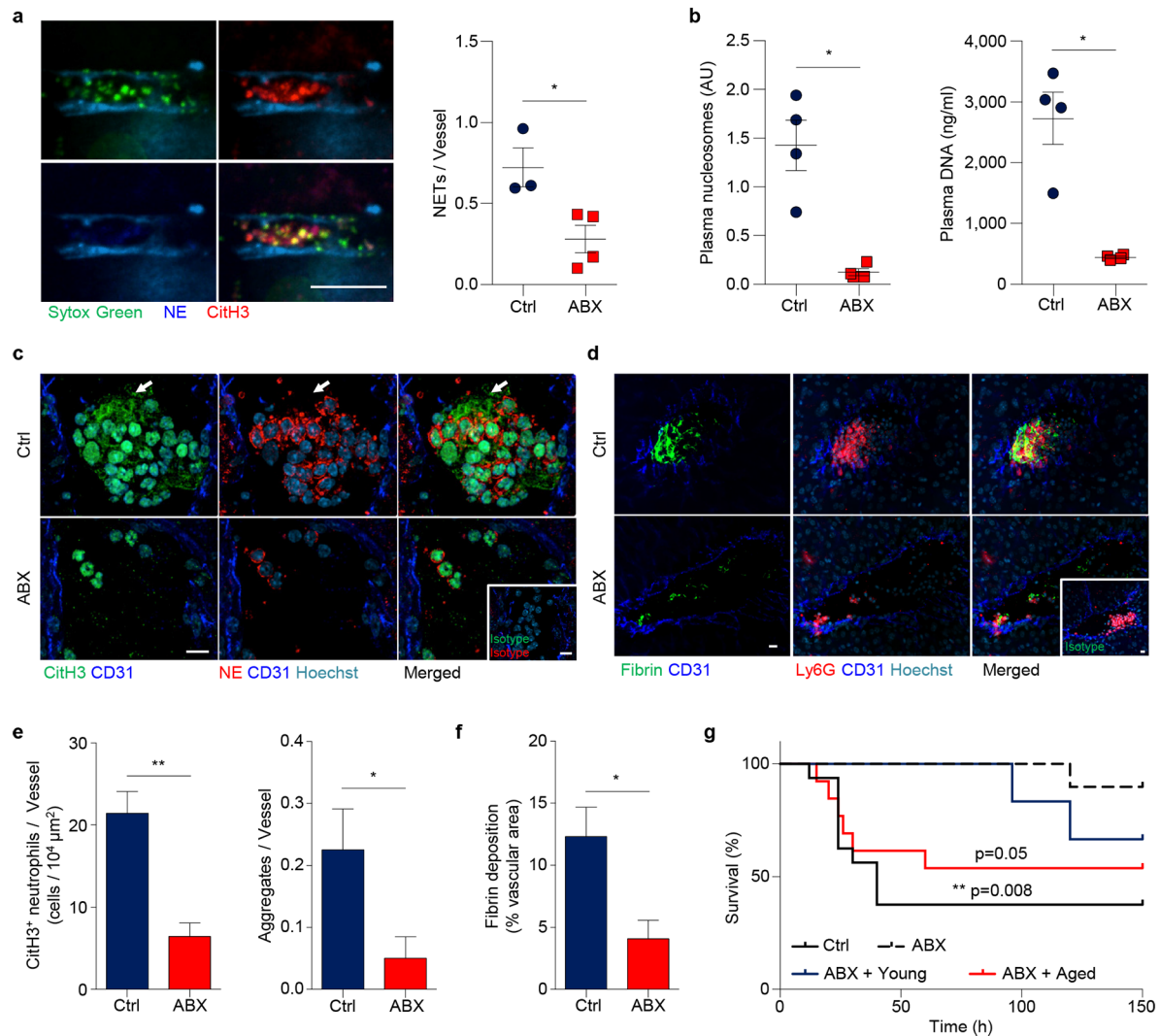
f, Percentages of aged neutrophils in wild-type, *Tlr4^{-/-}* and *Tlr2^{-/-}* mice ($n = 10, 10, 12$ mice, respectively). **g**, Flow cytometry analysis of aged neutrophils in wild-type and *Tnf^{-/-}* or *Csf2^{-/-}* mice. **h**, Percentages of wild-type and *LysM-cre/Myd88^{fl/fl}* or *Tlr4^{-/-}* or *Tlr2^{-/-}* neutrophils in total leukocyte population in chimeric mice ($n = 5$ mice). **i**, Percentages of wild-type and *LysM-cre/Myd88^{fl/fl}* or *Tlr4^{-/-}* or *Tlr2^{-/-}* neutrophils that capture more than eight beads in chimeric mice ($n = 5$ mice). Error bars, mean \pm s.e.m. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$, data representing two or more independent experiments analysed with unpaired Student's *t*-test (**a–f**) or paired Student's *t*-test (**h, i**).



Extended Data Figure 6 | Microbiota depletion inhibits NET formation.

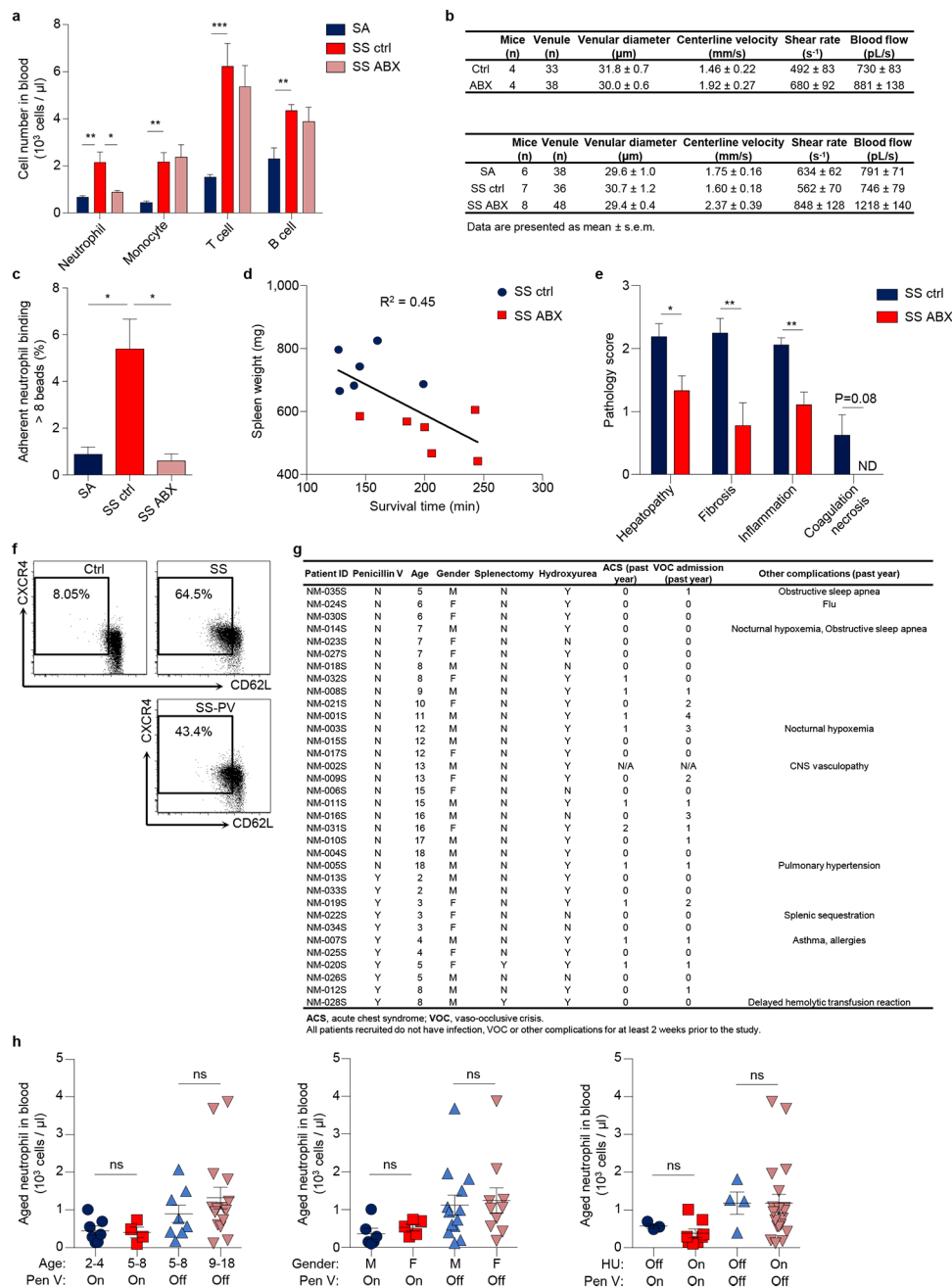
a, Flow cytometry analysis of aged neutrophils in isotype and anti-P/E-selectin antibody-treated mice ($n = 6, 5$ mice, respectively). **b**, ROS production of neutrophils from isotype and anti-P/E-selectin antibody-treated mice, as analysed by flow cytometry using dihydrorhodamine 123 (DHR-123; Isotype, $n = 10$; Abs (P/E), $n = 11$ mice). Grey lines, background fluorescence of neutrophils from both groups without LPS stimulation. ns, not significant. **c**, LPS-induced NET formation of neutrophils from control and antibiotics

(ABX)-treated mice, as analysed by immunofluorescence staining of DNA (sytox orange), neutrophil elastase (NE) and citrullinated histone 3 (CitH3). Inset, isotype control. Scale bars, 10 μ m. **d**, Quantification of NET formation of neutrophils from isotype and anti-P/E-selectin antibody-treated mice, or from control and ABX-treated mice (left, $n = 4$ (Isotype), 5 (Abs) mice; right, $n = 4$ mice). Error bars, mean \pm s.e.m. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$, data representing two or more independent experiments analysed with unpaired Student's t -test.



Extended Data Figure 7 | Microbiota depletion benefits endotoxin-induced septic shock. **a**, Representative images and quantification of *in vivo* NET formation in liver vasculature of control and antibiotics (ABX)-treated mice challenged with 30 mg kg⁻¹ LPS (*n* = 3, 4 mice, respectively). Scale bar, 10 μm. **b**, Quantification of NET biomarkers, plasma nucleosome and DNA, in septic control and ABX-treated animals (*n* = 4 mice). **c**, **d**, Representative images showing CitH3⁺ neutrophil aggregates (**c**) and fibrin deposition associated with neutrophil aggregates (**d**) in septic liver of control and ABX-treated mice. Arrows, diffusive CitH3 and neutrophil elastase (NE) proteins. Insets,

isotype controls. Scale bars, 10 μm. **e**, **f**, Numbers of CitH3⁺ neutrophils and neutrophil aggregates (**e**; left: *n* = 4 mice; right: *n* = 40 vessels from 4 mice) and quantification of fibrin deposition (**f**; *n* = 4 (Ctrl), 3 (ABX) mice) in septic liver of control and ABX-treated mice. **g**, Survival time of control, ABX-treated mice, and ABX-treated mice infused with 2 × 10⁶ aged or young neutrophils in septic shock induced by 30 mg kg⁻¹ LPS (*n* = 16, 10, 13, 6 mice, respectively). Error bars, mean ± s.e.m. **P* < 0.05, ***P* < 0.01, data representing two or more independent experiments analysed with unpaired Student's *t*-test (**a**, **e** (left), **f**), Mann-Whitney *U*-test (**b**, **e** (right)) or log-rank test (**g**).



Extended Data Figure 8 | Microbiota depletion affects disease progression in sickle-cell disease. **a**, Numbers of circulating leukocyte subsets in hemizygous control (SA), control SCD (SS Ctrl) and antibiotics-treated SCD (SS ABX) mice (SA: $n = 8$ mice; SS Ctrl: $n = 9$ mice; SS ABX: $n = 9$ mice). **b**, Haemodynamic parameters of mice analysed for neutrophil adhesion and integrin activation. **c**, Percentages of adherent neutrophils that capture more than eight beads in SA, SS Ctrl and SS ABX mice ($n = 4, 3, 3$ mice, respectively). **d**, Correlation between the survival times of SS control and SS ABX mice in acute vaso-occlusive crisis and their spleen weights. $R^2 = 0.45$. **e**, Scoring of liver damage, liver fibrosis, inflammation and necrosis in SS control and SS ABX

mice ($n = 8, 9$ mice, respectively). **f**, Flow cytometry analysis of aged neutrophils in healthy controls, SCD patients (SS), and SCD patients on penicillin V prophylaxis (SS-PV). **g**, Demographics of human subjects analysed for aged neutrophil numbers. ACS, acute chest syndrome; VOC, vaso-occlusive crisis. **h**, Aged neutrophil numbers in SCD patients grouped by age, gender, hydroxyurea (HU) and penicillin V (Pen V) treatment (Ctrl, $n = 9$ subjects; SS, $n = 23$ subjects; SS-PV, $n = 11$ subjects). Error bars, mean \pm s.e.m. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$, data representing two or more independent experiments analysed with unpaired Student's t -test (**a**, **c**, **h**) or Mann-Whitney U -test (**e**).

Extended Data Table 1 | Pathways selected for the analysis of neutrophil functions

Pathway	Gene set name
Innate immune system	REACTOME_INNATE_IMMUNE_SYSTEM
Local acute inflammatory response	BIOCARTA_LAIR_PATHWAY
Leukocyte adhesion and diapedesis	BIOCARTA_LYM_PATHWAY
MAPK signaling for integrins	REACTOME_P130CAS_LINKAGE_TO_MAPK_SIGNALING_FOR_INTEGRINS
Cell surface integrin pathway	PID_INTEGRIN_CS_PATHWAY
Fc-gamma receptor mediated phagocytosis	KEGG_FC_GAMMA_R_MEDIATED_PHAGOCYTOSIS
Lymphoid/Non-lymphoid interactions	REACTOME_IMMUNOREGULATORY_INTERACTIONS_BETWEEN_A_LYMPHOID_AND_A_NON_LYMPHOID_CELL
ROS pathway	BIOCARTA_FREE_PATHWAY
Peroxisome	KEGG_PEROXISOME
Cell/ECM interactions	REACTOME_CELL_EXTRACELLULAR_MATRIX_INTERACTIONS
Degradation of ECM	REACTOME_DEGRADATION_OF_THE_EXTRACELLULAR_MATRIX
Complement and coagulation cascades	KEGG_COMPLEMENT_AND_COAGULATION_CASCADES
TLR signaling pathway	KEGG_TOLL_LIKE_RECEPTOR_SIGNALING_PATHWAY
TLR endogenous pathway	PID_TOLL_ENDOGENOUS_PATHWAY
TLR4 signaling	REACTOME_ACTIVATED_TLR4_SIGNALLING
NLR signaling pathway	KEGG_NOD_LIKE_RECEPTOR_SIGNALING_PATHWAY
NLR pathway	REACTOME_NUCLEOTIDE_BINDING_DOMAIN_LEUCINE_RICH_REPEAT_CONTAINING_RECEPTOR_NLR_SIGNALING_PATHWAYS
NOD1/2 signaling pathway	REACTOME_NOD1_2_SIGNALING_PATHWAY
RLR pathway	KEGG_RIG_I_LIKE_RECEPTOR_SIGNALING_PATHWAY
RIG I mediated induction of IFN alpha	REACTOME_RIG_I_MDA5_MEDIATED_INDUCITION_OF_IFN_ALPHA_BETA_PATHWAYS
Activation of NFKB	REACTOME_ACTIVATION_OF_NF_KAPPAB_IN_B_CELLS
NFKB canonical pathway	PID_NFKAPPABCANONICALPATHWAY
NFKB atypical pathway	PID_NFKAPPABATYPICALPATHWAY
NFKB activation by RIP1	REACTOME_IKK_COMPLEX_RECRUITMENT_MEDIATED_BY_RIP1
NFKB activation through FADD/RIP1 pathway	REACTOME_NFKB_ACTIVATION_THROUGH_FADD_RIP1_PATHWAY_MEDIATED_BY_CASPASE_8_AND10
NFKB activation by TAK1	REACTOME_TAK1_ACTIVATES_NFKB_BY_PHOSPHORYLATION_AND_ACTIVATION_OF_IKKS_COMPLEX
NFKB and MAPK activation mediated by TLR4	REACTOME_NFKB_AND_MAP_KINASES_ACTIVATION_MEDIATED_BY_TLR4_SIGNALING_REPERTOIRE
NFKB activation induced by TRAF6	REACTOME_TRAF6_MEDIATED_INDUCITION_OF_NFKB_AND_MAP_KINASES_UPON_TLR7_8_OR_9_ACTIVATION
Chemokine signaling pathway	KEGG_CHEMOKINE_SIGNALING_PATHWAY
CXCR2 pathway	PID_IL8CXCR2_PATHWAY
Cytokine signaling pathway	REACTOME_CYTOKINE_SIGNALING_IN_IMMUNE_SYSTEM
Cytokines	BIOCARTA_CYTOKINE_PATHWAY
IL1 pathway	REACTOME_IL1_SIGNALING
IL2 receptor beta pathway	BIOCARTA_IL2RB_PATHWAY
IL4 pathway	ST_INTERLEUKIN_4_PATHWAY
IL6 pathway	REACTOME_IL_6_SIGNALING
IL12 pathway	PID_IL12_2PATHWAY
IL23 pathway	PID_IL23PATHWAY
IL27 pathway	PID_IL27PATHWAY
IFN-alpha signaling	REACTOME_REGULATION_OF_IFNA_SIGNALING
IFN-gamma pathway	REACTOME_INTERFERON_GAMMA_SIGNALING
TNFR1 pathway	BIOCARTA_TNFR1_PATHWAY
TNF/p75/NTR signaling	PID_P75NTRPATHWAY
TNFR2 pathway	BIOCARTA_TNFR2_PATHWAY
PPAR signaling pathway	KEGG_PPAR_SIGNALING_PATHWAY
p38/MAPK pathway	REACTOME_P38MAPK_EVENTS
ERK pathway	REACTOME_SIGNALLING_TO_ERKS
PIP3 signaling pathway	SIG_PIP3_SIGNALING_IN_CARDIAC_MYOCYTES
Ras signaling pathway	REACTOME_SIGNALLING_TO_RAS
RhoA pathway	PID_RHOA_PATHWAY
HDAC Class I pathway	PID_HDAC_CLASSI_PATHWAY
G-alpha I signaling pathway	REACTOME_G_ALPHA_I_SIGNALLING_EVENTS
NFAT pathway	PID_NFAT_TFPATHWAY
PDGF pathway	BIOCARTA_PDGF_PATHWAY
Rac1 pathway	PID_RAC1_PATHWAY
AKT pathway	BIOCARTA_AKT_PATHWAY
MAPK pathway	BIOCARTA_MAPK_PATHWAY
HIF pathway	BIOCARTA_HIF_PATHWAY
Purinergic receptor signaling pathway	REACTOME_NUCLEOTIDE_LIKE_PURINERGIC_RECEPTORS
Purinergic receptor P2Y signaling pathway	REACTOME_P2Y_RECEPTORS
Translation	REACTOME_TRANSLATION
Ribosome	KEGG_RIBOSOME
43S Ribosome	REACTOME_ACTIVATION_OF_THE_MRNA_UPON_BINDING_OF_THE_CAP_BINDING_COMPLEX_AND_EIFS_AND_SUBSEQUENT_BINDING_TO_43S
EIF pathway	BIOCARTA_EIF_PATHWAY
Protein export	KEGG_PROTEIN_EXPORT
Amino acid degradation	REACTOME_METABOLISM_OF_AMINO_ACIDS_AND_DERIVATIVES
Proteasome	KEGG_PROTEASOME
Ubiquitin mediated proteolysis	KEGG_UBIQUITIN_MEDIATED_PROTEOLYSIS
Protein degradation-Parkin pathway	BIOCARTA_PARKIN_PATHWAY
Cell death	BIOCARTA_DEATH_PATHWAY
Caspase cascades	SA_CASPASE_CASCADE
Cell death mitochondria pathway	BIOCARTA_MITOCHONDRIA_PATHWAY

Gene set name refers to Molecular Signatures Database v5.0 (Broad Institute).

Extended Data Table 2 | Gene set enrichment analysis of selected pathways in aged and activated neutrophils

Category	Pathway	Aged NES	Aged p-val	Activated NES	Activated p-val
Immune Functions	Innate immune system	1.2282217	0.10245901	2.1779234	0
	Local acute inflammatory response	1.1145728	0.3392857	1.8885438	0.001930502
	Leukocyte adhesion and diapedesis	1.5105772	0.06352941	1.6552799	0.025242718
	MAPK signaling for integrins	1.4485482	0.061440676	1.6242106	0.032128513
	Cell surface integrin pathway	1.4916337	0.0675	-1.2022164	0.24557522
	Fc-gamma receptor mediated phagocytosis	0.8664476	0.6876877	1.61418	0.008116883
	Lymphoid/Non-lymphoid interactions	0.8654324	0.6196172	1.5619569	0.033557046
	ROS pathway	-0.95446336	0.5321429	1.7730503	0.005565863
	Peroxisome	-1.203138	0.20919882	1.6480244	0.003395586
	Cell/ECM interactions	-0.7654254	0.7643979	1.772366	0.006048387
	Degradation of ECM	-1.4378743	0.095914744	1.5911175	0.028462999
	Complement and coagulation cascades	-1.5144768	0.070853464	1.6505035	0.030303031
	TLR signaling pathway	1.4612477	0.026706232	2.3468528	0
PRR and NFKB Signaling	TLR endogenous pathway	1.3665496	0.1056338	1.5889539	0.035714287
	TLR4 signaling	1.0985918	0.31333333	1.9638056	0
	NLR signaling pathway	1.4583313	0.05479452	2.5441191	0
	NLR pathway	1.4290915	0.07730673	2.1706495	0
	NOD1/2 signaling pathway	1.3088835	0.13197969	2.2286413	0
	RLR pathway	-0.8638851	0.6666667	2.1639493	0
	RIG I mediated induction of IFN alpha	-0.8472993	0.7259843	2.0400646	0
	Activation of NFKB	1.2855691	0.11396012	1.8813537	0.001718213
	NFKB canonical pathway	1.3068268	0.1495098	2.174795	0
	NFKB atypical pathway	0.8993623	0.6034483	1.9205661	0.001949318
	NFKB activation by RIP1	1.7887179	0.002192983	1.8208189	0.001930502
	NFKB activation through FADD/RIP1 pathway	1.8454865	0.008928572	1.6110392	0.033898305
	NFKB activation by TAK1	1.2061867	0.22624435	2.0932863	0
	NFKB and MAPK activation mediated by TLR4	1.1569227	0.22841226	1.8594357	0
	NFKB activation induced by TRAF6	1.0924268	0.27714285	1.6228119	0.010016695
Immune Functions	Chemokine signaling pathway	-0.7447356	0.8922652	1.433992	0.028616853
	CXCR2 pathway	-0.9767436	0.5085324	1.5124553	0.050583657
	Cytokine signaling pathway	-1.1273036	0.25921053	2.1093962	0
	Cytokines	-0.9237167	0.5686275	1.5562985	0.039325844
	IL1 pathway	1.6451061	0.02925532	2.078718	0
	IL2 receptor beta pathway	-2.0031374	0	1.2259008	0.2228261
	IL4 pathway	-1.312908	0.16470589	1.5741479	0.03169014
	IL6 pathway	-0.8152665	0.7010676	1.5760885	0.046747968
	IL12 pathway	-1.6290972	0.014240506	1.9707948	0.001675042
	IL23 pathway	-1.2130085	0.24518389	1.9758543	0
	IL27 pathway	-1.3348651	0.16555184	1.7636672	0.005464481
	IFN-alpha signaling	-0.8072854	0.68761224	1.6680452	0.014705882
	IFN-gamma pathway	1.4466043	0.087804876	2.066193	0
Signal Pathways	TNFR1 pathway	1.2722206	0.16010499	1.6132443	0.032432433
	TNF/p75/NTR signaling	1.3214921	0.08732394	1.726003	0.003231018
	TNFR2 pathway	-0.8648928	0.6243386	1.9728887	0
	PPAR signaling pathway	-0.8915187	0.6086956	2.0968451	0
	p38/MAPK pathway	-1.0720062	0.37918216	1.9020989	0
	ERK pathway	-1.1935399	0.24440895	1.4667499	0.04886562
	PIP3 signaling pathway	-1.4735942	0.05496183	1.5826265	0.022452503
	Ras signaling pathway	-1.6251277	0.02680067	1.4745739	0.07090909
	RhoA pathway	-1.6779591	0.017133957	1.1863663	0.24090122
	HDAC Class I pathway	-1.7304282	0.00608828	1.3537472	0.09764919
	G-alpha I signaling pathway	-0.8018528	0.78581977	1.4290625	0.03891709
	NFAT pathway	-1.5555012	0.039087947	-1.6272985	0.027272727
	PDGF pathway	-1.6846628	0.01983471	-1.7379107	0.014354067
	Rac1 pathway	0.64468014	0.91351354	1.5256119	0.033043478
Cellular Functions	AKT pathway	0.95451254	0.56	1.5437053	0.05009634
	MAPK pathway	1.1132134	0.25617284	1.637257	0.003289474
	HIF pathway	1.6175048	0.03117506	1.2415975	0.24675325
	Purinergic receptor signaling pathway	1.5389258	0.06153846	1.6881694	0.015968064
	Purinergic receptor P2Y signaling pathway	1.5777047	0.036876354	1.4960046	0.0503876
	Translation	1.4605397	0.028673835	-1.8208151	0
	Ribosome	1.1336436	0.23906706	-1.8144947	0
	43S Ribosome	1.1219336	0.28531855	-1.8783303	0
	EIF pathway	1.2240345	0.23933649	-1.5321815	0.05
	Protein export	1.5375326	0.06122449	-0.91629106	0.5726496
	Amino acid degradation	-1.5101556	0.018518519	1.3439316	0.05457464
	Proteasome	-1.3857354	0.06864274	1.4372953	0.06989247
	Ubiquitin mediated proteolysis	-0.926711	0.593361	1.7934686	0
	Protein degradation-Parkin pathway	-0.7039581	0.83516484	1.8072739	0.003731343
	Cell death	1.5957102	0.035128806	2.0140357	0
	Caspase cascades	1.5200465	0.065853655	1.9330438	0.001801802
	Cell death mitochondria pathway	1.6465071	0.029268293	1.6357895	0.016393442

NES, normalized enrichment score; p-val, nominal p value.

Loss of *Karma* transposon methylation underlies the mantled somaclonal variant of oil palm

Meilina Ong-Abdullah¹, Jared M. Ordway², Nan Jiang², Siew-Eng Ooi¹, Sau-Yee Kok¹, Norashikin Sarpan¹, Nuraziyan Azimi¹, Ahmad Tarmizi Hashim¹, Zamzuri Ishak¹, Samsul Kamal Rosli¹, Fadila Ahmad Malike¹, Nor Azwani Abu Bakar¹, Marhalil Marjuni¹, Norziha Abdullah¹, Zulkifli Yaakub¹, Mohd Din Amiruddin¹, Rajanaidu Nookiah¹, Rajinder Singh¹, Eng-Ti Leslie Low¹, Kuang-Lim Chan¹, Norazah Azizi¹, Steven W. Smith², Blaire Bacher², Muhammad A. Budiman², Andrew Van Brunt², Corey Wischmeyer², Melissa Beil², Michael Hogan^{2†}, Nathan Lakey², Chin-Ching Lim³, Xavier Arulandoo³, Choo-Kien Wong⁴, Chin-Nee Choo⁴, Wei-Chee Wong⁴, Yen-Yen Kwan⁵, Sharifah Shahrul Rabiah Syed Alwee⁵, Ravigadevi Sambanthamurthi¹ & Robert A. Martienssen⁶

Somaclonal variation arises in plants and animals when differentiated somatic cells are induced into a pluripotent state, but the resulting clones differ from each other and from their parents. In agriculture, somaclonal variation has hindered the micropropagation of elite hybrids and genetically modified crops, but the mechanism responsible remains unknown¹. The oil palm fruit ‘mantled’ abnormality is a somaclonal variant arising from tissue culture that drastically reduces yield, and has largely halted efforts to clone elite hybrids for oil production^{2–4}. Widely regarded as an epigenetic phenomenon⁵, ‘mantling’ has defied explanation, but here we identify the *MANTLED* locus using epigenome-wide association studies of the African oil palm *Elaeis guineensis*. DNA hypomethylation of a LINE retrotransposon related to rice *Karma*, in the intron of the homeotic gene *DEFICIENS*, is common

to all mantled clones and is associated with alternative splicing and premature termination. Dense methylation near the *Karma* splice site (termed the *Good Karma* epiallele) predicts normal fruit set, whereas hypomethylation (the *Bad Karma* epiallele) predicts homeotic transformation, parthenocarpy and marked loss of yield. Loss of *Karma* methylation and of small RNA in tissue culture contributes to the origin of mantled, while restoration in spontaneous revertants accounts for non-Mendelian inheritance. The ability to predict and cull mantling at the plantlet stage will facilitate the introduction of higher performing clones and optimize environmentally sensitive land resources.

The African oil palm (*E. guineensis*) is the most efficient oil-bearing crop, but demand for edible oils and biofuels, combined with sustainability concerns over dwindling rainforest reserves, has led to intense

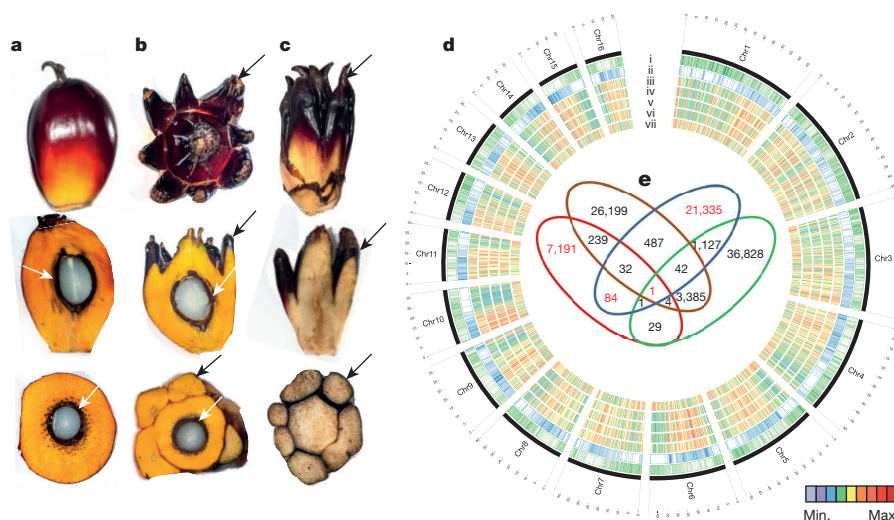


Figure 1 | Epigenome-wide association study (EWAS). **a–c**, Normal (a), fertile mantled (b) and parthenocarpic mantled (c) fruit shown as whole fruit (top), longitudinal sectioned (middle) and cross sectioned (bottom). Black arrows denote pseudocarpetals; white arrows denote kernel. **d**, Circos plot of oil palm chromosomes. Track order: gene density (i); repeat density (ii); cytosine methylation density (whole-genome bisulfite sequencing) in ortet (iii); cytosine methylation densities (microarray) of ortet (iv), normal ramet (v) and mantled ramet (vi); differential cytosine methylation of normal minus mantled ramets (vii). Heatmaps represent average cytosine methylation densities in ~300-kb

windows independent of sequence context. **e**, Venn diagram of microarray features differentially methylated between leaves from mantled and normal ramets ($P < 0.05$, two-sided Student *t*-test, Methods). Each set represents clonal lineages of given genotypes obtained from one source: source A (red, 15 mantled, 15 normal), source B (brown, 6 mantled, 14 normal), source C (blue, 14 mantled, 15 normal) and source D (green, 8 mantled, 10 normal). Red numbers indicate subsets including one of the four microarray features mapping to the *Karma* LINE element.

¹Malaysian Palm Oil Board, 6, Persiaran Institusi, Bandar Baru Bangi, 43000 Kajang, Selangor, Malaysia. ²Orion Genomics, 4041 Forest Park Avenue, St Louis, Missouri 63108, USA. ³United Plantations Berhad, Jendarata Estate, 36009 Teluk Intan, Perak, Malaysia. ⁴Applied Agricultural Resources Sdn Bhd, No. 11, Jalan Teknologi 3/6, Taman Sains Selangor 1, 47810 Kota Damansara, Petaling Jaya, Selangor, Malaysia. ⁵FELDA Global Ventures R&D Sdn Bhd, c/o FELDA Biotechnology Centre, PT 23417, Lengkuu Teknologi, 71760 Bandar Enstek, Negeri Sembilan, Malaysia. ⁶Howard Hughes Medical Institute-Gordon and Betty Moore Foundation, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York 11724, USA. †Present address: Thermo Fisher Scientific, 110 Miller Avenue, Ann Arbor, Michigan 48104, USA.

pressure to improve oil palm yield. Introduction of the *tenera* hybrid (*dura* × *pisifera*) increased oil yield by up to 30%, leveraging the *SHELL* gene that confers single gene heterosis^{6,7}. Clones (ramets) of individual high-yielding *tenera* hybrid palms (ortets) provide a powerful shortcut to yield enhancement, with an additional 20–30% improvement⁶. Micropropagation through cell culture of immature apex leaf tissue (the ‘heart of palm’), and plantlet regeneration on hormone-supplemented media (Methods), yields tens of thousands of genetically identical clonal palms. Unfortunately, shortly after the procedure was established, Tan Yap Pau of United Plantations, Malaysia, first noted a high frequency of homeotic floral phenotypes known as ‘mantling’ among clonal ramets (T.Y.P., personal communication). Subsequently, Corley *et al.*² documented the occurrence of mantled palms after prolonged periods in culture. In mantled palms, staminodes of pistillate flowers and stamens of staminate flowers develop as pseudocarpels⁸, often resulting in sterile parthenocarpic flowers with abortive fruit and very low oil yields (Fig. 1a–c and Extended Data Fig. 1). Pollination of mantled palms gave rise to variable numbers of mantled progeny, resembling rare naturally mantled variants known as *poissoni*, or *diwakkawakka* fruit forms^{3,4}. The trait is non-Mendelian and sometimes reverts to normal⁹ and so has long been considered epigenetic⁵, with an overall decrease in DNA methylation found in mantled ramets^{5,10}. The homeotic transformations observed in mantled palms resemble defects in B-function MADS-box genes, suggesting strong candidates for epigenetic modification⁸. However, decades of research into candidate retroelements^{11,12} and candidate homeotic genes^{8,12,13} have failed to identify epigenetic changes consistently found in somaclonal mantled palms.

We performed a genome-wide, unbiased, DNA methylation analysis (an epigenome-wide association study; EWAS) in search of loci epigenetically associated with the mantled phenotype, using a DNA microarray based on the *E. guineensis* (*pisifera*) reference genome¹⁴ (Methods). DNA methylation density was measured in 1–2-kilobase (kb) intervals surrounding each feature by DNA-methylation-dependent comparative microarray hybridization¹⁵ and statistical analyses (Methods). Genome-wide DNA methylation maps were constructed from parthenocarpic mantled ($n = 43$) or normal ($n = 54$) ramets, as well as ortets from which these ramets were derived ($n = 10$). These maps strongly resembled those constructed by whole-genome bisulfite sequencing in sample palms (Fig. 1d), demonstrating reproducibility.

At genome-wide resolution, the landscape of DNA methylation was remarkably consistent between ortets and ramets (Fig. 1d), with highest methylation within repetitive sequences¹⁴. However, thousands of loci were differentially methylated (Fig. 1d), most of which (~90%) were hypomethylated in mantled, consistent with previously reported reduced 5mC content^{5,10}. Most hypomethylated loci (~75%) were transposons and repeats, while less frequent hypermethylated loci included genic sequences (Extended Data Fig. 2), resembling cell cultures of *Arabidopsis*¹⁶. Fifteen independent somaclonal lineages obtained from four independent sources were used to maximize genotypic diversity, and significant differentially methylated regions (DMRs) between normal and fully mantled samples were first identified within each source population (Methods). Results were then compared between populations (Fig. 1e). Although tens of thousands of DMRs were detected between mantled and normal clones in each population, 99.9% of these were exclusive to either one (94.4%) or two

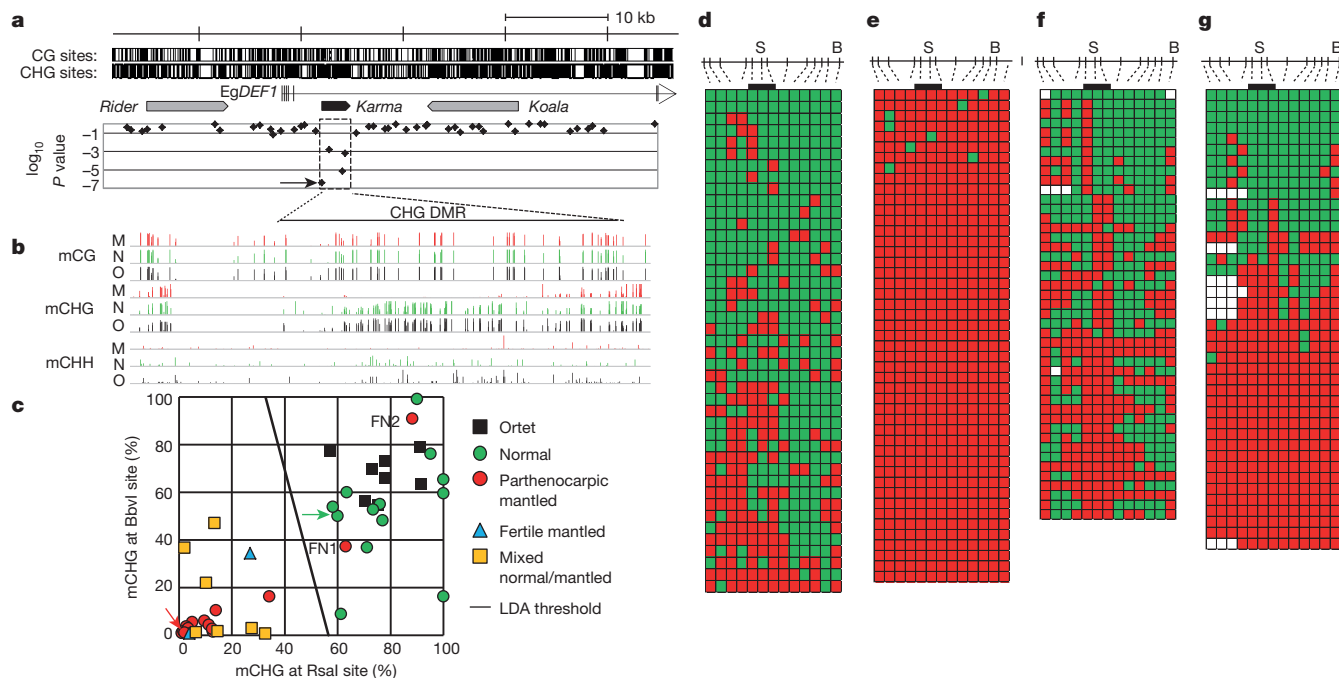


Figure 2 | Hypomethylation of *Karma* is associated with the mantled phenotype. **a**, Microarray feature data plotted below a map of the *EgDEF1* gene (vertical ticks, exons; horizontal line, introns; arrow, direction of transcription) including locations of *Rider*, *Karma* (dashed box) and *Koala* retrotransposons. CG and CHG sites are shown at the top. \log_{10} P values (54 normal versus 43 parthenocarpic mantled ramets) are plotted (two-sided Student's t -test). Arrow in P value plot denotes feature detected as hypomethylated in mantled ramets from all four sources (Fig. 1e). **b**, Genome-wide bisulfite sequencing of leaf samples from ortet (O, black, $n = 5$), normal ramets (N, green, $n = 5$) and parthenocarpic mantled ramets (M, red, $n = 5$). Mean methylation density per cytosine is plotted on a 0–100% scale for each cytosine context and sample type. CHG DMR, differentially CHG methylated region corresponding to *Karma*. **c**, CHG methylation monitored in 86 additional ortets, mantled and

normal ramet leaf samples by restriction enzyme digestion and qPCR (Methods). Linear discriminant analysis was performed between normal ($n = 21$) and mantled ($n = 28$) samples with BbvI and RsaI restriction sites. FN1 and FN2, two false-negative mantled samples. Green and red arrows denote normal and mantled control samples, respectively. A similar analysis was performed on remaining normal ($n = 14$) and mantled ($n = 23$) samples with ScrFI restriction sites (Extended Data Fig. 4d). **d–g**, *Karma* bisulfite sequencing maps (antisense strand) of normal control (**d**), mantled control (**e**), FN1 (**f**) and FN2 (**g**). Thirteen CHG sites are shown to scale above. ‘S’ denotes CHG at the *Karma* splice acceptor site (CAG/CTG); ‘B’ denotes the BbvI site. Bar, CHGs within the common microarray feature (Fig. 1e). Methylated and unmethylated CHG sites are indicated by green and red boxes, respectively. Open boxes denote low-quality base calls. Each row represents an individual Sanger DNA sequencing read.

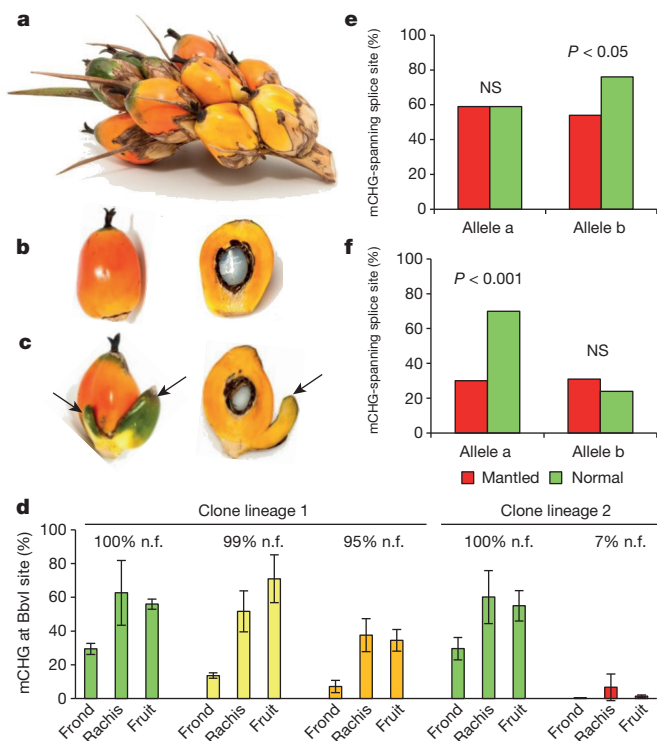


Figure 3 | *Karma* methylation in revertant palms. **a–c**, Spikelet from a revertant ramet (**a**) including normal (**b**) and fertile (**c**) mantled fruit with one or two pseudocarps (arrows). **d**, Density of CHG methylation (percentage mCHG) at the BbvI site (Methods) in ramets yielding 100% normal fruit (n.f.) (green), revertant ramets yielding 99% (yellow) or 95% (orange) normal fruit and a mosaic ramet yielding 7% (red) normal fruit per bunch. Error bars denote s.d. (biological replicates of fronds ($n = 4$), rachis sections ($n = 8$) or fruit ($n = 2$)). **e, f**, Percentage mCHG for the three CHG sites found in the unique common microarray feature in normal (green) and subtly mantled (red) fruit from revertant ramets yielding 99% (e) or 95% (f) normal fruit per bunch (two-tailed Fisher's exact test; NS, not significant). Alleles were analysed separately based on a heterozygous single nucleotide polymorphism (SNP) within the bisulfite sequencing amplicon.

(5.5%) of the four populations, indicating considerable genotypic variation in epigenetic response to tissue culture. A single microarray feature detected differential methylation between normal and mantled clones in all four populations (Fig. 1e). This feature lies within the ~35 kb intron 5 of *EgDEF1* (Fig. 2a), the oil palm orthologue of the B-class MADS-box transcription factor genes, *Antirrhinum majus* *DEFICIENS* (*DEF*) and *Arabidopsis* *APETALA3* (*AP3*)^{8,12,13}.

Elaeis guineensis *DEF1* spans ~40 kb on chromosome 12 (Fig. 2a). A *Ty1/copia* *Rider* retrotransposon lies upstream, while a *Ty3/gypsy* retrotransposon, *Koala*, is located within intron 5. Consistent with important earlier work¹², no DNA methylation difference within these retrotransposons was found in mantled clones across several populations (Fig. 2a and Extended Data Fig. 3). However, a third previously unreported repetitive element lies within intron 5, and has homology to rice *Karma* LINE elements. *Karma* is activated in rice embryogenic tissue culture, but only transposes in regenerated plants as transgenerational DNA hypomethylation of the element persists¹⁷. The 3.2-kb oil palm *Karma* element is flanked by a 13-base-pair (bp) target site duplication (TTCAAAATGATGA) and includes a defective reverse transcriptase open reading frame (ORF2) preceded by a splice acceptor (‘^’) and followed by a polyadenylation signal, resembling truncated *Karma* elements in rice^{17,18} (Supplementary Fig. 1). The unique microarray feature, which consistently detected hypomethylation in mantled clones, serendipitously includes the predicted splice acceptor site (GAACAG^ATGC). All three additional microarray features mapping within the *Karma* element also detected significant

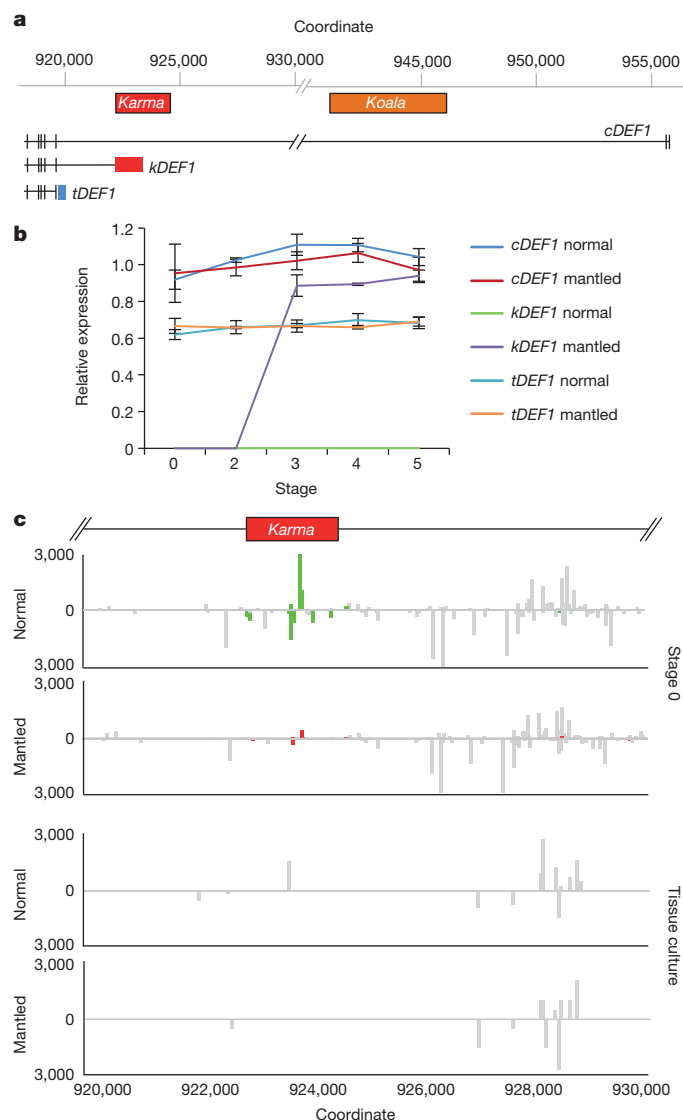


Figure 4 | Alternative splicing and loss of 24-nucleotide siRNA. **a**, *EgDEF1* transcripts assembled from transcriptome sequencing (data not shown) and RT-PCR (Methods). Black boxes denote exons; blue box denotes intron 5 sequence included in the *tDEF1* transcript. Coordinates relative to the reference *pisifera* oil palm genome¹⁴. **b**, Quantitative RT-PCR of *cDEF1*, *tDEF1* and *kDEF1* transcripts in shoot apices (stage 0) and in early (stage 2) to late (stage 5) female inflorescences from normal and parthenocarpic mantled ramets. Error bars, s.d. between three replicate assays of three replicate tissue samples per phenotype, per stage. Expression relative to an endogenous reference gene is shown (Methods). **c**, 24-nucleotide siRNA accumulation in shoot apices (stage 0) from normal ($n = 5$) and parthenocarpic mantled ($n = 7$) ramets, and from second passage apical leaf tissue cultures re-cloned from normal ($n = 2$) or mantled ($n = 1$) ramets (Methods). Values expressed as fragments per kilobase per million mapped reads (FPKM). Bars above (sense) and below (antisense) the line indicate mapped normalized 24-nucleotide siRNAs that are not significantly different in abundance in normal and mantled (grey) or significantly differentially expressed in normal (green) relative to mantled (red) ($P < 0.05$, Student's *t*-test, two tailed, assuming equal variance).

hypomethylation in mantled clones (Fig. 2a and Extended Data Figs 3 and 4a–c).

To verify *Karma* hypomethylation, sample trios comprising genetically identical ortet, parthenocarpic mantled and normal ramets from five independent clonal lineages were subjected to whole-genome bisulfite sequencing (Methods). CG methylation was unchanged across the *EgDEF1* locus, but *Karma* CHG methylation ($H = A$, C

or T) was markedly reduced in mantled clones, revealing a DMR covering ~70 CHG sites. CHH methylation was much lower and only subtly reduced (Fig. 2b). To validate differential CHG methylation in unrelated clonal palms, quantitative PCR (qPCR) assays were used to quantify CHG methylation at BbvI and RsaI restriction sites within the DMR (Methods, Fig. 2c and Extended Data Fig. 4b) in a panel of 49 palms from 21 clonal lineages and 4 independent sources: 8 ortets and 13 normal clones, 19 parthenocarpic mantled clones, 2 fertile mantled clones and 7 partially revertant clones yielding bunches with both mantled and normal fruit. Linear discriminant analysis provided 93% sensitivity and 100% specificity for detection of mantling (Fig. 2c). Fronds from all seven of the revertant palms were scored as mantled, consistent with the observation that normal bunches arose late in development². Similar results were obtained in 37 polymorphic palms using alternative restriction sites (Extended Data Fig. 4d). The two false-negative mantled palms (Fig. 2c) were further analysed by bisulfite sequencing of a region spanning the *Karma* splice acceptor site. While normal clones had dense CHG methylation, and mantled controls had lost all CHG methylation (Fig. 2d, e), the false-negative mantled samples lost CHG methylation near the splice acceptor site (Fig. 2f, g and Extended Data Fig. 4e), which was therefore sufficient to predict the mantled phenotype. Because of their strong predictive properties, we named the mantled hyper- and hypomethylated epialleles *Good Karma* and *Bad Karma*, respectively.

Two lineages of revertant palms had mixed bunches with both normal and mantled fruit⁹, resembling epialleles in maize regulated by transposons¹⁹. The first lineage included two revertant ramets with 99% and 95% normal fruit per bunch, respectively, in which abnormal fruit had only one or two small pseudocarps (Fig. 3a–c). A second lineage included a mosaic ramet with only 7% normal fruit. In all three ramets, CHG methylation at the BbvI site was low in fronds (Fig. 3d), consistent with other revertants (Fig. 2c). However, methylation was restored in fruit from the two revertant ramets, but not from the mantled mosaic ramet (Fig. 3d–f). As with similar epialleles in maize, *Linaria* and other plants^{19–21}, reversion of the abnormal phenotype accompanied by restoration of DNA methylation is strong evidence that *Karma* hypomethylation is the cause of the mantled phenotype. Differential methylation between individual mantled and normal fruit was not observed, however, probably reflecting non-cell autonomy of the B-class homeotic phenotype (Fig. 3a–d), also observed in *Antirrhinum* and *Arabidopsis*²². Bisulfite sequencing from normal and mantled fruit (Extended Data Fig. 5) revealed hyper- and hypomethylated reads at the splice acceptor site (Fig. 3e, f), suggesting that these fruit were indeed mosaic for hyper- and hypomethylated cells. In one mosaic palm, direct evidence for mosaicism was obtained from different samples of the same vegetative frond (Extended Data Fig. 6).

DNA methylation near splice acceptor sites affects alternative splicing, although the mechanism remains unclear²³. To assess alternative splicing, *EgDEF1* transcript models were built on the basis of transcriptome sequencing (data not shown), and were validated by reverse transcription PCR (RT–PCR; Methods). As previously reported¹², two forms of *EgDEF1* transcripts were found in both normal and mantled inflorescences: the full-length *EgDEF1* transcript (*cDEF1*), and a prematurely terminated transcript including exons 1–5 and 221 bp of intron 5 (*tDEF1*)¹² (Fig. 4a and Extended Data Fig. 4c). However, we identified a third alternative transcript in mantled female inflorescences (Fig. 4a). This novel transcript (*kDEF1*) was spliced from the donor site of intron 5 to the proximal *Karma* acceptor site, and is predicted to encode a truncated *EgDEF1* peptide that terminates within the K domain of the MADS-box protein, and has a unique carboxy-terminal sequence (Extended Data Fig. 7a).

Expression of the three transcripts (Fig. 4b) was assessed in shoot apical meristem (stage 0) and four stages of immature female inflorescence development (Methods and Extended Data Fig. 7b–f). *kDEF1* expression was notably restricted to stages 3–5 of mantled (but not normal) female inflorescence, strongly suggestive of a role in the

mantled phenotype. By contrast, *cDEF1* was detected at only slightly (although significantly) lower levels in mantled female inflorescences¹², while *tDEF1* was unchanged (Fig. 4b). Previously reported differences in the timing of *tDEF1* and *cDEF1* transcription within each stage¹² may be the consequences of *kDEF1* transcription (Methods).

In plants, 24-nucleotide small interfering RNAs (siRNAs) guide non-CG methylation, and we identified a cluster of antisense 24-nucleotide *Karma* siRNAs in shoot apical meristem (stage 0), which were reduced or absent in mantled (Fig. 4c) and in later stages of normal inflorescence (Extended Data Fig. 8 and Methods). In polyembryonic tissue cultures derived from normal and abnormal clonal palms (Methods), small RNA (sRNA) underwent a switch from 24 to 21 nucleotides (Extended Data Fig. 9), resembling cell cultures and callus of *Arabidopsis*¹⁶. *Karma* methylation (Extended Data Fig. 10) and 24-nucleotide siRNA (Fig. 4c and Extended Data Fig. 8) were reduced in normal cultures between two and seven passages, and lost in abnormal cultures, suggesting a model for the origin of mantled: if meristems are the source of 24-nucleotide siRNA²⁴ (Fig. 4c), then leaf cells detached from the meristem would progressively lose 24-nucleotide siRNA and non-CG DNA methylation over time in culture. Antisense sRNA might influence exon trapping, as it does in humans^{25,26}, or else splicing might be associated with changes in chromatin, for example histone H3 Lys4 methylation¹⁸. *Good Karma* would only be restored during shoot generation if DNA methylation and small RNA were not entirely lost during tissue culture, or potentially if siRNAs were artificially applied during the tissue culture process. The 24-nucleotide siRNA could also facilitate non-Mendelian segregation⁴⁹, which resembles paramutation (the interconversion of heterozygous epialleles) in some respects^{11,23,27}.

Despite its importance, mantling has been the elusive target of molecular genetic investigation for the past three decades. We have demonstrated that the mantled trait is a consequence of epigenetic modification of the *Karma* transposable element within the B-class MADS-box *EgDEF1* gene, which we have named *MANTLED*. B-class function in stamen identity is conserved in monocots, but similar to other monocot paleoAP3 genes, *EgDEF1* overexpression fails to cause homeotic conversion in *Arabidopsis*⁸, because of a diverged C-terminal exon²⁸. Nonetheless, like AP3 and *DEF*^{29,30}, the oil palm *MANTLED* gene is expressed in inner perianth and stamen primordia as the perianth initiates (stage 2), followed by stamen and stamenoïd primordia (stage 3)⁸. The appearance of *kDEF1* transcripts during this transition (Fig. 4b) suggests that *kDEF1* functions at this crucial window to induce the mantled phenotype.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 30 April; accepted 10 August 2015.

Published online 9 September 2015.

1. Stroud, H. *et al.* Plants regenerated from tissue culture contain stable epigenome changes in rice. *Elife* **2**, e00354 (2013).
2. Corley, R. H. V. in *CRC Handbook of Fruit Set and Development* (ed. Monselise, S. P.) 253–259 (CRC Press, 1986).
3. Zeven, A. C. The 'mantled' oil palm (*Elaeis guineensis* Jacq.). *J. W. Afr. Inst. Oil Palm Res.* **5**, 31–33 (1973).
4. Mgbeze, G. C. & Iserhienrhien, A. Somaclonal variation associated with oil palm (*Elaeis guineensis* Jacq.) clonal propagation: a review. *Afr. J. Biotechnol.* **13**, 989–997 (2014).
5. Jaligot, E., Rival, A., Beule, T., Dussert, S. & Verdeil, J. L. Somaclonal variation in oil palm (*Elaeis guineensis* Jacq.): the DNA methylation hypothesis. *Plant Cell Rep.* **19**, 684–690 (2000).
6. Corley, R. H. V. & Law, I. H. in *Plantation Management for the 21st Century* (ed. Pushparajah, E.) 279–289 (Incorp. Soc. Planters, 1997).
7. Singh, R. *et al.* The oil palm *SHELL* gene controls oil yield and encodes a homologue of SEEDSTICK. *Nature* **500**, 340–344 (2013).
8. Adam, H. *et al.* Functional characterization of MADS box genes involved in the determination of oil palm flower structure. *J. Exp. Bot.* **58**, 1245–1259 (2007).
9. Rao, V. & Donough, C. R. Preliminary evidence of a genetic cause for the floral abnormalities in some oil palm ramets. *Elaeis* **2**, 199–207 (1990).
10. Matthes, M., Singh, R., Cheah, S. C. & Karp, A. Variation in oil palm (*Elaeis guineensis* Jacq.) tissue culture-derived regenerants revealed by AFLPs with methylation-sensitive enzymes. *Theor. Appl. Genet.* **102**, 971–979 (2001).

11. Kubis, S. E., Castilho, A. M., Vershinin, A. V. & Heslop-Harrison, J. S. Retroelements, transposons and methylation status in the genome of oil palm (*Elaeis guineensis*) and the relationship to somaclonal variation. *Plant Mol. Biol.* **52**, 69–79 (2003).
12. Jaligot, E. *et al.* DNA methylation and expression of the *EgDEF1* gene and neighboring retrotransposons in *mantled* somaclonal variants of oil palm. *PLoS ONE* **9**, e91896 (2014).
13. Syed Alwee, S. *et al.* Characterization of oil palm MADS box genes in relation to the mantled flower abnormality. *Plant Cell Tissue Organ Cult.* **85**, 331–344 (2006).
14. Singh, R. *et al.* Oil palm genome sequence reveals divergence of interfertile species in Old and New worlds. *Nature* **500**, 335–339 (2013).
15. Lippman, Z. *et al.* Role of transposable elements in heterochromatin and epigenetic control. *Nature* **430**, 471–476 (2004).
16. Tanurdzic, M. *et al.* Epigenomic consequences of immortalized plant cell suspension culture. *PLoS Biol.* **6**, e302 (2008).
17. Komatsu, M., Shimamoto, K. & Kyoizuka, J. Two-step regulation and continuous retrotransposition of the rice LINE-type retrotransposon Karma. *Plant Cell* **15**, 1934–1944 (2003).
18. Cui, X. *et al.* Control of transposon activity by a histone H3K4 demethylase in rice. *Proc. Natl Acad. Sci. USA* **110**, 1953–1958 (2013).
19. Martienssen, R., Barkan, A., Taylor, W. C. & Freeling, M. Somatic heritable switches in the DNA modification of Mu transposable elements monitored with a suppressible mutant in maize. *Genes Dev.* **4**, 331–343 (1990).
20. Cubas, P., Vincent, C. & Coen, E. An epigenetic mutation responsible for natural variation in floral symmetry. *Nature* **401**, 157–161 (1999).
21. Saze, H. & Kakutani, T. Heritable epigenetic mutation of a transposon-flanked *Arabidopsis* gene due to lack of the chromatin-remodeling factor DDM1. *EMBO J.* **26**, 3641–3652 (2007).
22. Perbal, M. C., Haughn, G., Saedler, H. & Schwarz-Sommer, Z. Non-cell-autonomous function of the *Antirrhinum* floral homeotic proteins *DEFICIENS* and *GLOBOSA* is exerted by their polar cell-to-cell trafficking. *Development* **122**, 3433–3441 (1996).
23. Regulski, M. *et al.* The maize methylome influences mRNA splice sites and reveals widespread paramutation-like switches guided by small RNA. *Genome Res.* **23**, 1651–1662 (2013).
24. Baubec, T., Finke, A., Mittelsten Scheid, O. & Pecinka, A. Meristem-specific expression of epigenetic regulators safeguards transposon silencing in *Arabidopsis*. *EMBO Rep.* **15**, 446–452 (2014).
25. Taniguchi-Ikeda, M. *et al.* Pathogenic exon-trapping by SVA retrotransposon and rescue in Fukuyama muscular dystrophy. *Nature* **478**, 127–131 (2011).
26. Alló, M. *et al.* Control of alternative splicing through siRNA-mediated transcriptional gene silencing. *Nature Struct. Mol. Biol.* **16**, 717–724 (2009).
27. Arteaga-Vazquez, M. *et al.* RNA-mediated trans-communication can establish paramutation at the *b1* locus in maize. *Proc. Natl Acad. Sci. USA* **107**, 12986–12991 (2010).
28. Lamb, R. S. & Irish, V. F. Functional divergence within the *APETALA3/PISTILLATA* floral homeotic gene lineages. *Proc. Natl Acad. Sci. USA* **100**, 6558–6563 (2003).
29. Sommer, H. *et al.* *Deficiens*, a homeotic gene involved in the control of flower morphogenesis in *Antirrhinum majus*: the protein shows homology to transcription factors. *EMBO J.* **9**, 605–613 (1990).
30. Jack, T., Brockman, L. L. & Meyerowitz, E. M. The homeotic gene *APETALA3* of *Arabidopsis thaliana* encodes a MADS box and is expressed in petals and stamens. *Cell* **68**, 683–697 (1992).

Supplementary Information is available in the online version of the paper.

Acknowledgements We acknowledge the contributions of staff members of the Breeding and Tissue Culture Unit at MPOB for creating the valuable clonal lines, and for their extensive data collection and sampling efforts. We thank Genomics Unit at MPOB for conducting DNA fingerprinting to verify clonal lines. We thank The McDonnell Genome Institute at Washington University for genomic bisulfite sequencing and transcriptome sequencing support, and MGen for microarray hybridizations. At Orion Genomics, we thank N. Sander, J. Reed, J. Brune, K. Soe, J. McDonald, C. Brown and B. Dove for technical support, and M.-F. Wu and M. Sachdeva for assistance with the manuscript and additional informatics support. We would also like to thank T. Dalmay for recommendations on sRNA library construction. We appreciate the constant support of the Director-General of MPOB, Datuk Dr. Yuen-May Choo, and the Ministry of Plantation Industries and Commodities, Malaysia.

Author Contributions M.O.-A. led the work on the *MANTLED* marker/gene. M.O.-A., R.Si., E.-T.L.L. and R.Sa. conceptualized the research programme. M.O.-A., R.Si., E.-T.L.L., R.N., N.L., S.W.S., J.M.O., R.Sa. and R.A.M. developed the overall strategy, designed experiments and coordinated the project. A.T.H., Z.I. and S.K.R. performed tissue culture on selected ortets and field-planted the ramets. Field data collection and fruit bunch census were conducted at various research stations by F.A.M., N.A.A.B., M.M., N.A., Z.Y. and M.D.A. M.O.-A., C.-C.L., X.A., C.-N.C., W.-C.W., S.S.R.S.A. and Y.-Y.K. identified samples for discovery and validation panels. M.O.-A., C.-C.L. and X.A. identified materials used in the mosaic experiments. M.O.-A., S.-E.O., S.-Y.K., N.S. and N.A. conducted laboratory experiments, histological staging of inflorescences and data analyses. N.J. and S.W.S. performed microarray analyses. B.B. and M.A.B. prepared fractions for microarray hybridizations. B.B. designed and analysed qPCR experiments. B.B. and M.B. performed qPCR assays. A.V.B. designed and analysed clone-based bisulfite sequencing experiments, and A.V.B. and M.B. performed bisulfite sequencing assays. M.B. designed and performed qRT-PCR experiments. C.W. and J.M.O. analysed transcriptome data. K.-L.C., N.A., S.W.S., M.H., C.W. and A.V.B. provided bioinformatics support. M.O.-A., R.Si., E.-T.L.L., R.N., N.L., S.W.S., J.M.O., R.Sa. and R.A.M. prepared and revised the manuscript.

Author Information Microarray data have been deposited in the NCBI Gene Expression Omnibus (GEO) and are accessible through GEO Series accession number GSE68410. Small RNA sequence data from the region of interest have been deposited in the NCBI Sequence Read Archive (SRA) database under the accession numbers SAMN03569290–SAMN03569351. Whole-genome bisulfite sequence data have been deposited in the NCBI SRA under accession numbers SAMN03569063–SAMN03569077. The cDNA sequence of the *kDEF1* transcript has been deposited in GenBank under accession number KR347486. Reprints and permissions information is available at www.nature.com/reprints. The authors declare competing financial interests: details are available in the online version of the paper. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to R.A.M. (martiens@csh.edu) or R.Sa. (raviga@mpob.gov.my).

METHODS

Samples and tissue culture. Leaf tissue used in the EWAS discovery and validation panels was sampled from oil palm clones (known as ramets) derived from *tenera* mother palms (known as ortets) from various genetic backgrounds. The female parents of the ortets were predominantly of a Deli *dura* background while the *pisifera* male parents were derived from La Me, Yangambi, AVROS and Binga genotypes. These parental lines make up the major genetic backgrounds of the oil palm populations in Malaysia. These palms were collected from the Malaysian Palm Oil Board (MPOB), United Plantations Berhad, FELDA Global Ventures R&D Sdn Bhd and Applied Agricultural Resources Sdn Bhd (Extended Data Fig. 3). The ramets were derived from explants excised from non-chlorophyllous leaves of their respective ortets cultured on hormone-supplemented media^{31–34}. The oil palm tissue culture process involves callus initiation and polyembryoid generation followed by shoot and root initiation³⁵. In large scale production, the tissue culture process takes 48–52 months to complete. Once established, ramets undergo acclimatization in the nursery for 3–4 months, are moved to the field nursery for another 8–9 months, and finally are field planted. At every stage of the tissue culture process, off-types are culled. Flower and fruit bunch census are taken at the onset of flowering (2–3 years after field planting) and in subsequent years. Normal and mantled ramets were identified based on the census data collected.

Recloned tissue culture materials were generated using the standard protocol described above. Normal and abnormal ramets from identical genetic backgrounds were subjected to the tissue culture process and sub-sampling was carried out at the polyembryogenic stage, namely at subculture passage two (SC2) and seven (SC7), representing short and prolonged exposure in culture on hormone-supplemented media, respectively.

Female inflorescence samples used in staging the developmental phase were obtained from a total of 31 clonal palms with ages ranging from 3 to 10 years after field-planting. As a means to categorize the developmental phases, these inflorescence samples were histologically analysed and classified as stage 0, shoot apical meristem; stage 2, initiation of perianth organs; stage 3, development of perianth organs and initiation of reproductive organs; stage 4, development of reproductive organs; stage 5, fully formed reproductive organs according to ref. 36.

***E. guineensis* genome microarray design.** The microarray design was based on the *E. guineensis* genome¹⁴ and contains more than 1 million 60-base probes. Probes were selected from non-overlapping 1.5-kb windows across all scaffolds of the P1 *pisifera* genome build, choosing the 60-mer with the lowest composite 15-mer frequency count within the genome. When compared to the publically available EG5 genome assembly¹⁴, 860,861 probes from the microarray matched at 100% stringency (81,194 probes falling on exons and the remaining covering intronic and intergenic regions of the genome). The microarrays were manufactured by Roche NimbleGen using the HX1 platform.

DNA methylation-dependent fractionation. Genomic DNA (60 µg) was mechanically sheared to 1–4-kb fragments. Sheared DNA was divided into four equal portions, two of which were digested with 10 U µg⁻¹ McrBC (New England Biolabs) under manufacturer's recommended conditions and the other two were mock-treated without adding enzyme. After digestion, DNAs were treated with proteinase K (50 mg ml⁻¹) for 1 h at 50 °C and then precipitated with ethanol under standard conditions. Resuspended DNAs were resolved by agarose gel electrophoresis, and DNA in the 1–4-kb size range was excised from gels and extracted. McrBC requires that two methylated half sites (RmC, where R = A or G) lie within 40–3,000 bp of each other, and cutting occurs in the proximity of one of the half sites³⁷. Because 1–4-kb fragments are treated with McrBC, and undigested fragments are isolated, hybridization of a microarray probe complementary to sequence distant from the methylation site results in a 'wingspan' effect in which probes are able to detect DNA methylation from a distance up to ~1.5 kb³⁸. For each fraction, 200 ng was used for cyanogen dye labelling (Cy3 or Cy5). For each sample, four microarrays were hybridized in a duplicated dye swap design to differentially labelled untreated and DNA methylation depleted fractions. Sample size was chosen to allow several clonal lineages including both normal and parthenocarpic samples from each of four independent sources, but no statistical methods were used to determine sample size.

Microarray data processing, normalization and statistical analysis. Among the ~860,000 microarray probes matching the oil palm genome with 100% sequence stringency, a subset of ~460,000 uniquely mapped probes was selected to reduce noise from non-specific hybridizations. For data processing, corrections on spatial non-uniformity of fluorescent signal intensities, done separately for the Cy3 and Cy5 dye channels, were made. Data were then normalized by background subtraction using negative control probes, followed by scaling. After normalization, median log signal of the control probes for each dye of each array was set to zero. MAD (median absolute deviation) log signal of all the probes on an array was used as a constant for a given treatment (untreated or DNA methylation depleted). DNA methylation was measured as the sample average (over the two pairs of

dye-swapped technical replicates) of normalized log₂ ratios of untreated over methylation-depleted DNA. Statistical analysis was first conducted within samples derived from each source independently. Within each group, a two-sided *t*-test was performed between normal and mantled phenotypes. Using a cut-off of *P* = 0.05, one probe that was significantly differentially methylated in all the four groups was identified. To confirm this finding with a different statistical approach, quantile normalization of methylation measurements (the sample average of normalized log ratios of untreated over methylation depleted DNA) was performed on all samples together, and then a *t*-test on all normal versus all mantled samples was conducted. This process identified the same probe that was found in the initial analysis, as well as an additional three immediately neighbouring probes. The experiments were not randomized, and investigators were not blinded to allocation during experiments and outcome assessment.

Code availability. Computer code for microarray data processing and normalization is available for download at http://www.oriongenomics.com/files/methylscope_processing.tar.

qPCR DNA methylation assays. Primer pairs were designed to amplify two *Karma* element regions as diagrammed in Extended Data Fig. 4a, b. A 633-bp amplicon included methylation-sensitive restriction sites containing CHG positions 188 bp (BbvI) and 375 bp (ScrFI) downstream of the *Karma* splice site CHG. A 632-bp amplicon that amplified a region near the centre of the *Karma* element included a RsaI methylation-sensitive restriction site. Primer pairs were confirmed to amplify a single band of the correct size by agarose gel electrophoresis. For qPCR DNA methylation assays, 100 ng of genomic DNA was digested with 10 U µg⁻¹ of the indicated restriction enzyme under standard conditions. An equal amount of genomic DNA was mock-treated in a reaction lacking enzyme. Digestion reactions were incubated at 37 °C for 16 h. qPCR was carried out using 10 ng each of the mock-treated and enzyme-digested samples in 1 × Roche SYBR Green Master Mix on a Roche LC480 instrument. qPCR amplifications were performed in duplicate. For each duplicate mock/digested amplification pair, the ΔC_t value was calculated as the digested C_t minus the mock C_t and duplicated ΔC_t values were averaged. DNA methylation density was calculated as percentage dense methylation = $2^{(-\Delta C_t(\text{digested} - \text{mock}))}$. Samples were genotyped by restriction digestion of PCR amplicons with either BbvI or RsaI to confirm that all samples used for DNA methylation validation included intact restriction sites on both alleles. Enzyme digestions were quality controlled by performing qPCR assays monitoring three independent invariantly unmethylated endogenous genomic loci and one invariantly methylated endogenous genomic locus. All quality control passed digestions reported < 5% methylation of the unmethylated controls and > 95% methylation of the methylated control. Primer sequences are available on request.

Whole-genome bisulfite sequencing. Genomic DNA (1 µg) from each of 15 mature leaf samples (5 lineage trios of ortet, normal ramet and mantled ramet) was used to construct TruSeq fragment libraries (Illumina), and up to 500 ng of adapted library molecules were bisulfite converted using the EZ DNA Methylation-Lightning Kit (Zymo Research). Each library was sequenced in one lane of a HiSeq 2000 flow cell to generate paired 100-bp reads. Reads were mapped to an *in silico* bisulfite converted reference *E. guineensis* (*pisifera*) genome. For each cytosine context (CG, CHG or CHH), the number of mapped reads corresponding to unconverted cytosines relative to the total number of reads including the particular base was used to calculate the percent methylation at each cytosine position.

Clone based bisulfite sequencing. Because all cytosines are potential sites for DNA methylation in plants, and because whole-genome bisulfite sequencing demonstrated that CG methylation is maintained at high levels in both normal and mantled ramets, bisulfite sequencing amplicon primers were designed to include CG dinucleotides, but exclude CHG and CHH trinucleotides. Within primer sequences, CG dinucleotides were assumed to be methylated. The amplicon (amplified from the antisense strand) contained 13 CHG sites, including the CHG site at the *Karma* splice acceptor site. Then 2 µg of each sample was bisulfite-converted as described for whole-genome bisulfite sequencing. In total, 30 ng converted DNA was used for PCR amplification in 1 × HiFi Hotstart Uracil+ Ready Mix (Kappa). Amplicons were cloned using the TOPO TA Cloning Kit (Invitrogen) following A-tailing by Klenow treatment. For each sample, 48 white colonies were individually picked, propagated and plasmid DNA extracted. Plasmid inserts were PCR-amplified and Sanger-sequenced (ABI 3730) using vector-specific primers. Sequencing was performed on 48 clones per sample, and reads from plasmids not including the amplicon insert are not shown. Sequences were base called in CONSED and methylation densities at each CHG site were calculated. Where possible, heterozygous non-cytosine SNPs were scored so that each allele could be analysed independently. In cases where a polymorphism changed a CHG site to either a CG or CHH site, the non-CHG variant was not included in calculations of CHG methylation. Because CHH methylation was

determined to be consistently very low in both normal and mantled ramets, conversion of CHH sites within the amplicon was used to control for bisulfite conversion rates. All samples analysed displayed <4% methylation of CHH sites, demonstrating that bisulfite conversion was >96% complete in all samples. Primer sequences are available on request.

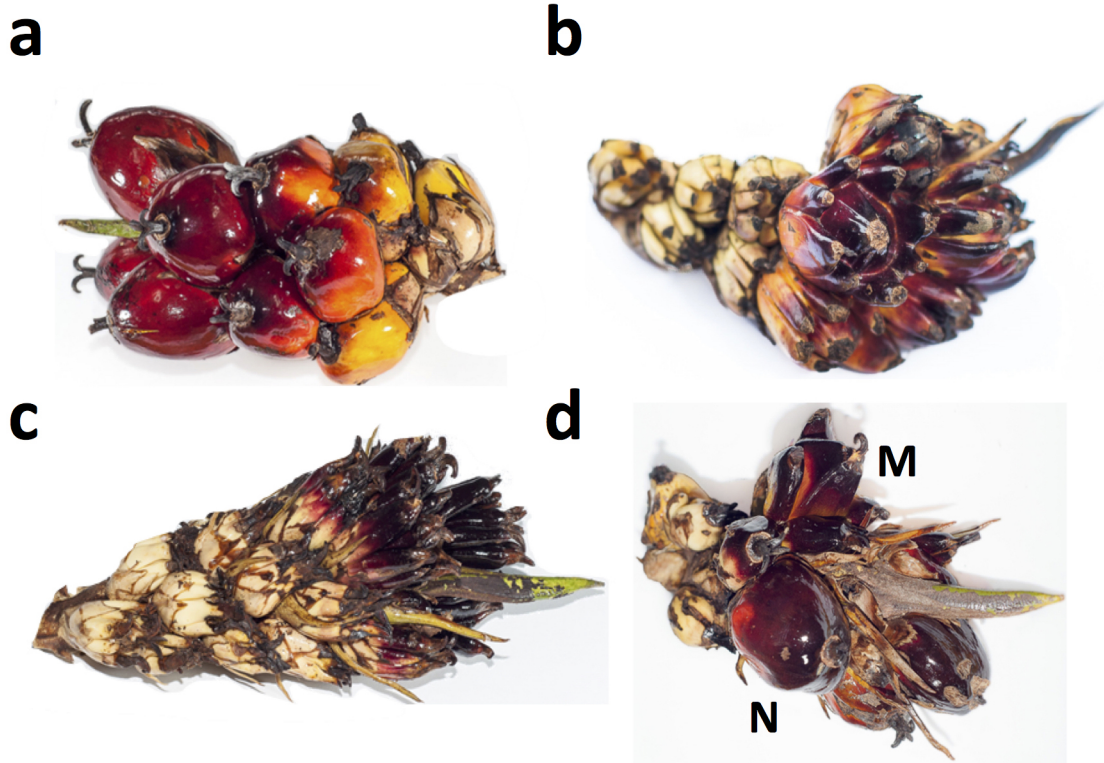
mRNA and siRNA sequencing. Transcriptome sequencing was performed on shoot apex, early inflorescence (<2 cm) and late stage inflorescence (three normal and three mantled female inflorescence biological replicates each). About 2–3 µg total RNA was used to construct individually barcoded Illumina TruSeq stranded libraries. Libraries were pooled in sets of four, and each pool was sequenced in one lane of a HiSeq 2000 flow cell to generate paired 100-bp reads. sRNA fractions of female shoot apex tissue at stage 0 and female inflorescence tissue at stages 2, 3, 4 and 5 (7 mantled and 5 normal biological replicates at stage 0, 6 mantled and 8 normal biological replicates each at stages 2 and 3, 7 mantled and 5 normal biological replicates at stage 4, and 5 mantled and 4 normal biological replicates at stage 5), as well as second-passage tissue cultures recloned from mantled ($n = 1$) or normal ($n = 2$) ramets, were used to construct Illumina TruSeq sRNA libraries and sequenced following the same strategy as mRNA sequencing. mRNA sequencing data was used to construct gene models for all observed *EgDEF1* alternative transcripts. siRNA reads mapping to the genomic scaffold including *EgDEF1* were identified and normalized as fragments per 1,000 reads mapped to the genome (FPKM). FPKM values for each 24-mer were compared between biological replicates of normal and mantled samples by a two-tailed Student's *t*-test, assuming equal variance.

qRT-PCR assays. To specifically quantify *cDEF1* expression, a forward primer spanning the junction of exons 1 and 2 was used with a reverse primer within exon 7 (Extended Data Fig. 7b–d). The same forward primer was used with a reverse primer including intron 5 sequence to specifically quantify *tDEF1* expression. Finally, a forward primer spanning the junction of exons 4 and 5 was used with a reverse primer within *Karma*, downstream of the exon 5/*Karma* splice junction, to specifically quantify *kDEF1* expression. Assays were optimized using normal and mantled late stage inflorescence total RNA, and cDNAs were Sanger sequenced to confirm the identity of the amplicons. Standard curves generated from serially diluted cDNAs were generated for each primer pair, as well as for two internal oil palm reference gene assays³⁹ (Extended Data Fig. 7e, f). Gene expression was quantified in developing inflorescence stages 0, 2, 3, 4 and 5. All first-strand cDNA reverse transcription reactions were performed from 1 µg total RNA using a cocktail of reverse primers specific to *EgDEF1* exons 6 and 7, as well as 3' regions of *Karma*. For each stage, three technical replicates were performed for each of the three biological replicates per phenotype, per stage. qRT-PCR reac-

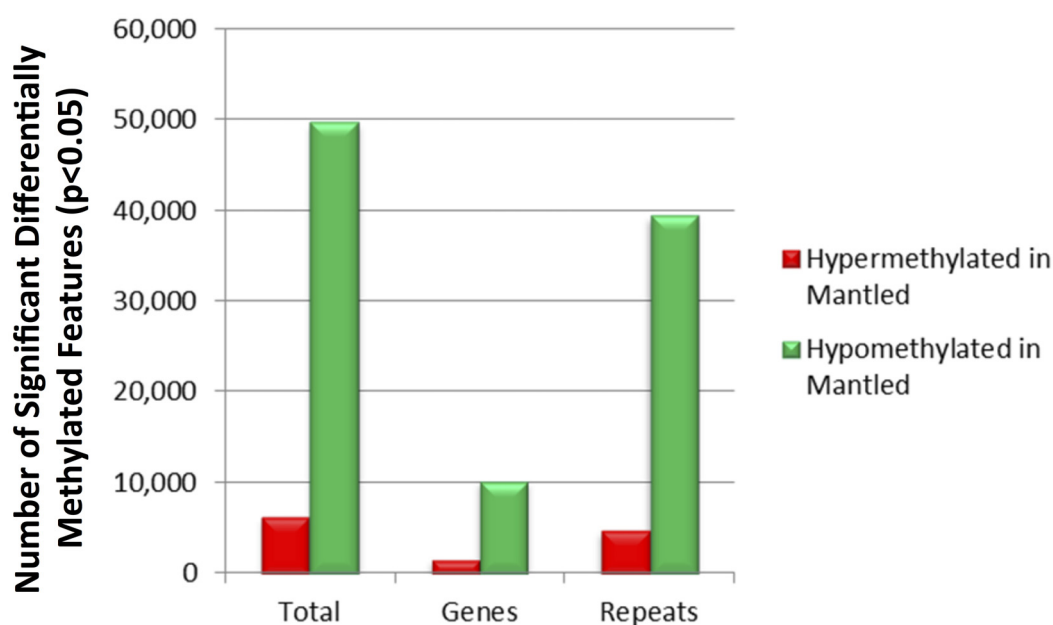
tions were performed using 1 µl first-strand cDNA in 1× Roche SYBR Master Mix on a Roche LC480 instrument. Cycle thresholds above 33 cycles were not included in calculations, and detectable expression was calculated only for samples in which expression was detected in at least 2 of 3 technical replicates. Expression levels were quantified by extrapolation from the standard curve for each assay, and expression levels relative to the reference gene were calculated. Primer sequences are provided in Extended Data Fig. 7d.

We found only a subtle decrease in expression of *cDEF1* in female mantled relative to normal inflorescence at stage 3 (Fig. 4b). A more significant decrease was previously reported⁸ but was not detected consistently⁴⁰. More recently, an increase in the ratio of *tDEF1* to *tDEF1* + *cDEF1* in samples from early and late time points at each stage of inflorescence development was reported¹². However, when absolute values of (*tDEF1*/*tDEF1* + *cDEF1*) from early and late samples from any given stage are pooled (see figure 6b and Supplementary figure 8b in ref. 12) then only modest increases in expression are observed in mantled samples in agreement with our results (Fig. 4b).

31. Jones, L. H. Propagation of clonal oil palms by tissue culture. *Planter* **50**, 374–381 (1974).
32. Rabechault, H. & Martin, J.-P. Multiplication végétative du palmier à huile (*Elaeis guineensis* Jacq.) à l'aide de cultures de tissus foliaires. *C. R. Acad. Sci. Paris. Ser. D* **283**, 1735–1737 (1976).
33. Durand-Gasselin, T., Le Guen, V., Konan, K. & Duval, Y. Oil palm (*Elaeis guineensis* Jacq.) plantations in Côte-d'Ivoire obtained through *in vitro* culture. First results. *Oléagineux* **45**, 10–11 (1990).
34. Kushairi, A. *et al.* Current status of oil palm tissue culture in Malaysia. In *Proceedings of the Clonal and Quality Replanting Material Workshop. Towards Increasing the Annual National Productivity by One Tonne FFB/ha/year 3–14* (Malaysian Palm Oil Board, 2006).
35. Rohani, O. *et al.* in *Advances in Oil Palm Research* (eds Yusof, B., Jalani, B. S. & Chan, K. W.) 238–283 (Malaysian Palm Oil Board, 2000).
36. Adam, H. *et al.* Determination of flower structure in *Elaeis guineensis*: do palms use the same homeotic genes as other species? *Ann. Bot.* **100**, 1–12 (2007).
37. Stewart, F. J. & Raleigh, E. A. Dependence of McrBC cleavage on distance between recognition elements. *Biol. Chem.* **379**, 611–616 (1998).
38. Ordway, J. M. *et al.* Comprehensive DNA methylation profiling in a human cancer genome identifies novel epigenetic targets. *Carcinogenesis* **27**, 2409–2423 (2006).
39. Chan, P. L. *et al.* Evaluation of reference genes for quantitative real-time PCR in oil palm elite planting materials propagated by tissue culture. *PLoS ONE* **9**, e99774 (2014).
40. Jaligot, E. *et al.* Epigenetic imbalance and the floral developmental abnormality of the *in vitro*-regenerated oil palm *Elaeis guineensis*. *Ann. Bot.* **108**, 1453–1462 (2011).

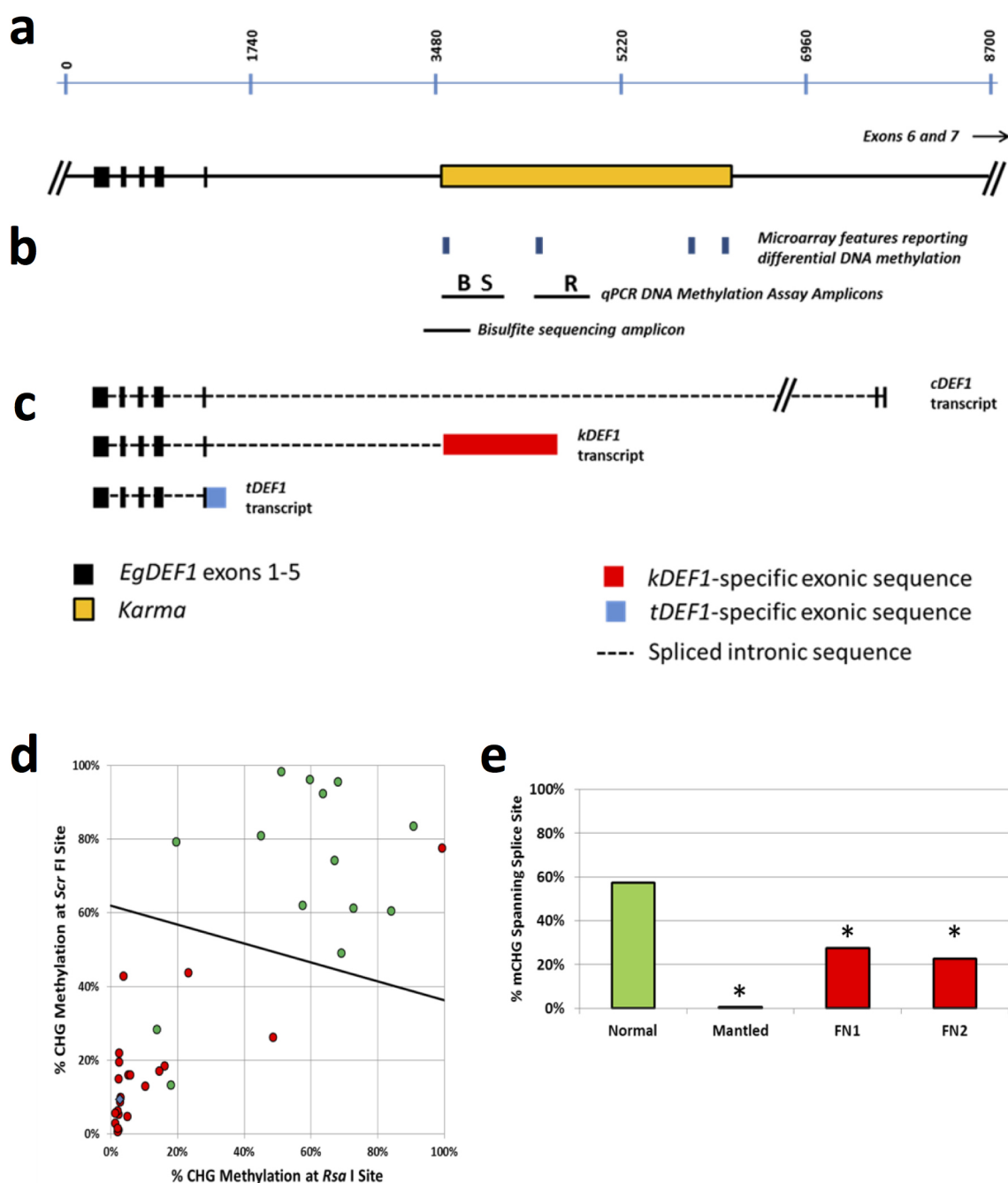


Extended Data Figure 1 | Spikelets from clonal palms of different fruit form phenotypes. **a**, Spikelets from a normal ramet. **b**, Spikelet from a fertile mantled ramet. **c**, Spikelet from a parthenocarpic mantled ramet. **d**, Spikelet from a revertant ramet displaying both normal (N) and mantled (M) fruits in the same spikelet.



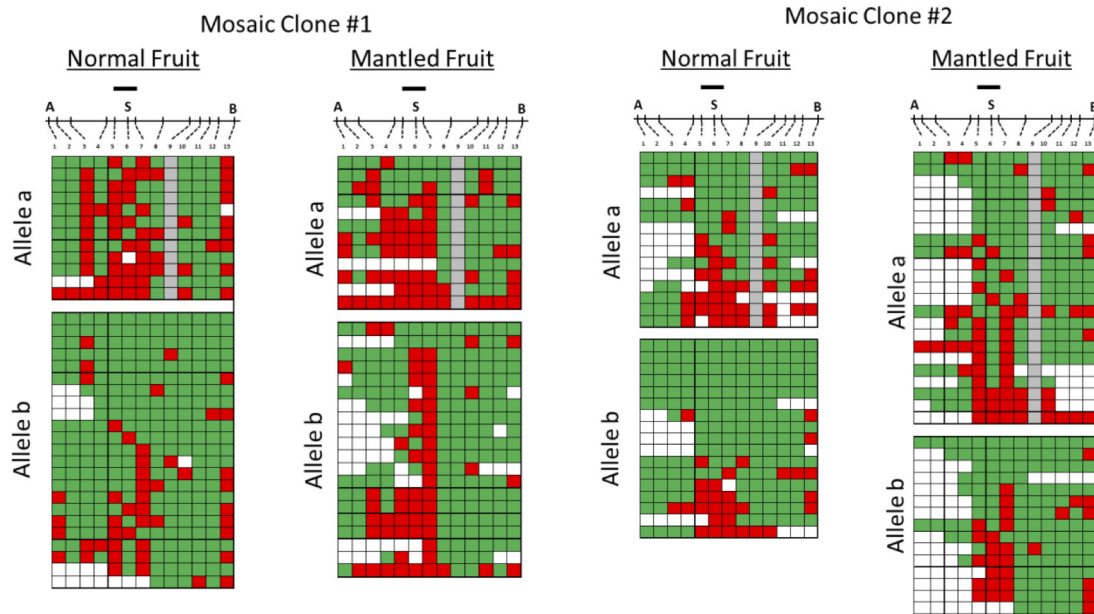
Extended Data Figure 2 | Annotation of genome-wide differentially methylated loci. Sequences of microarray features reporting significant differential DNA methylation between fully normal and fully mantled leaf DNA samples of one or more clonal lineages ($P < 0.05$, two-sided Student's *t*-test, Methods) were mapped to the reference *E. guineensis pisifera* genome¹⁴. Numbers of biological replicates per clonal lineage are provided in Extended Data Fig. 3. Features were assigned to gene and repeat classes

according to annotations of genomic elements mapped within 3 kb of the microarray feature sequence, as this is the distance at which McrBC is capable of monitoring DNA methylation density. The repeat class includes all repetitive sequences, including transposons and *pisifera*-specific repetitive sequences¹⁴. Features mapping within 3 kb of both a gene and a repeat were assigned to both classes. The number of features reporting hypermethylation (red) and hypomethylation (green) are plotted.



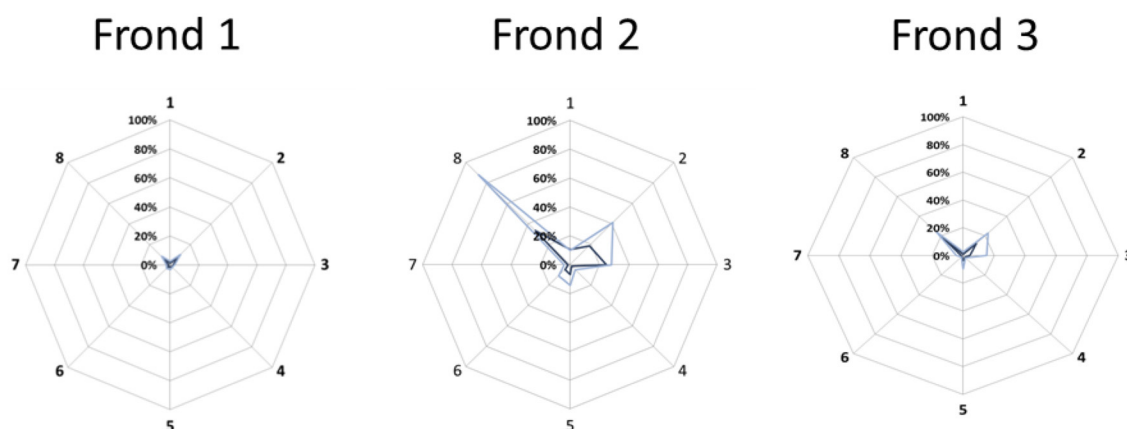
Extended Data Figure 4 | DNA methylation assays and supporting DNA methylation data. **a**, Diagram of the *EgDEF1* gene including the *Karma* element within intron 5 (orange box). Black boxes represent exons and the horizontal line represents introns. Scale bar is in base pair units. **b**, Blue tick marks represent the relative positions of the four microarray features reporting significant hypomethylation of mantled clones in all source lineages. The left-most feature includes the *Karma* splice acceptor site. Horizontal lines labelled B (BbvI) and R (RsaI) indicate the relative positions of amplicons used for qPCR-based CHG methylation assays. The BbvI amplicon also includes a ScrFI site (S) used in **d**. The relative position of the bisulfite sequencing amplicon used to determine *Karma* splice site CHG methylation is shown below the qPCR amplicons. **c**, Diagrams of the three alternatively spliced *EgDEF1* transcripts. Black boxes represent exons included in each transcript. The dotted lines represent intronic sequences spliced out of the mature mRNA transcripts. The red box represents *Karma* element sequence spliced to

EgDEF1 exon 5 in the *kDEF1* transcript. The blue box represents *EgDEF1* intron 5 sequence included in the *tDEF1* transcript that does not use the exon 5 splice donor site. **d**, In addition to adult leaf samples analysed by BbvI and RsaI qPCR assays (Fig. 2c), 37 samples were found to have a SNP in the BbvI site and were therefore analysed by ScrFI and RsaI qPCR assays (Methods and Extended Data Fig. 4b). Linear discriminant analysis was performed between normal ($n = 14$) and mantled ($n = 22$ parthenocarpic mantled; $n = 1$ fertile mantled) samples. Combining these results with those shown in Fig. 2c, sensitivity and specificity for detection of mantling are each 94%. **e**, Bisulfite sequencing of controls, FN1 and FN2 (Fig. 2c). mCHG density was calculated for the three CHG sites covered by the unique common microarray feature (Figs 1e and 2d–g). FN1, FN2 and the mantled control were significantly hypomethylated relative to the normal control ($*P < 0.0001$, two-tailed Fisher's exact test).



Extended Data Figure 5 | Clone based bisulfite sequencing maps of normal and mantled phenotype fruits from epigenetic mosaics. The heatmap format is as described in Fig. 2d–g. Grey boxes indicate a site in which a SNP on allele a results in a CHG to CHH site conversion. Mosaic clone 1 represents a revertant clone yielding 95% normal fruit. Mosaic clone 2 represents a

revertant clone yielding 99% normal fruit. Alleles were analysed independently based on a SNP not affecting a potentially methylated base. Statistical analyses of methylation at the three CHG sites spanning the *Karma* splice site are shown in Fig. 3e, f.



Extended Data Figure 6 | CHG methylation in rachis sectors of an oil palm yielding 7% normal fruit (clone lineage 2 in Fig. 3d). Rachis of three successive fronds was dissected into 8 equal sectors. DNA methylation in each sector per frond was measured by BbvI and ScrFI assays, as described in Methods. Average DNA methylation density measurements of three technical replicates per frond, per sector, per assay are plotted on a radial graph representing the 8 rachis sections around the palm trunk (ScrFI assay, light blue; BbvI assay dark blue). Sector numbering was ratcheted for frond 2 versus

1, and frond 3 versus 2 based on the R^2 best fit of CHG methylation density around the circumference of the palm to correct for out-of-register numbering of rachis sectors between successive fronds (data not shown). Consistent with the fact that this oil palm yields only 7% normal phenotype fruit, most DNA methylation measurements are consistent with the mantled phenotype. However, sectors 8 and 2 display gains of CHG methylation in rachis sectors of all three fronds, and reach or approach normal levels in sectors 8 and 2 of frond 2, thus demonstrating mosaicism directly.

a

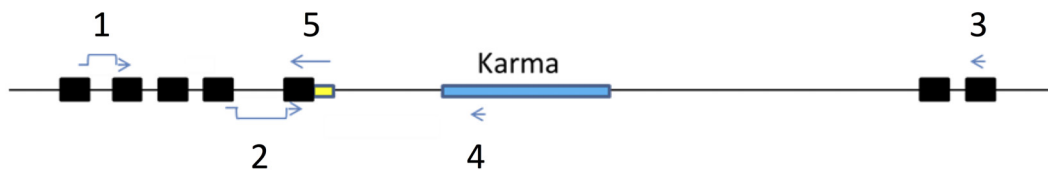
1 : MGRGKIEIKKIENPTNRQVTYSKRRTGIMKKAKELTVLCDAEVSLIMFSSTGKFSEYCSPLSDT : 64 kDEF1
 1 : MGRGKIEIKKIENPTNRQVTYSKRRTGIMKKAKELTVLCDAEVSLIMFSSTGKFSEYCSPLSDT : 64 cDEF1

65 : KTIFDRYQQVSGINLWSAQYEKMQNTLNHLREINQNLRRERQRMGEDLDSLGIHELRLGLEQNL : 128
 65 : KTIFDRYQQVSGINLWSAQYEKMQNTLNHLREINQNLRRERQRMGEDLDSLGIHELRLGLEQNL : 128

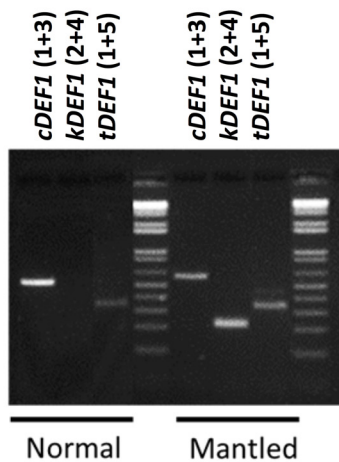
129 : DEALKVVHRHKYHVITTQTDYKKK**MHLKSALDHLLK** : 156 +12
 129 : DEALKVVHRHKYHVITTQTDYKKKLKNSNEAHKNLLHELEMKDEHPVYGFVDDDDPSNYAGALA : 192

193 : LANGASHMYAFRVQPSQPNLHRMGFGSHDLRLA* : 226

b



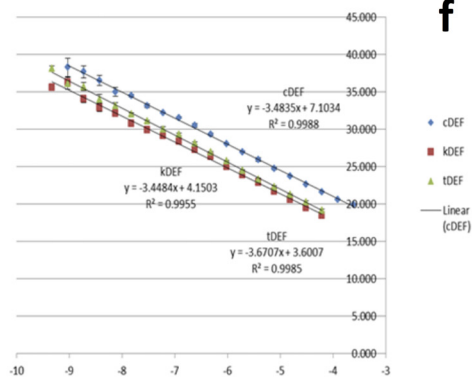
c



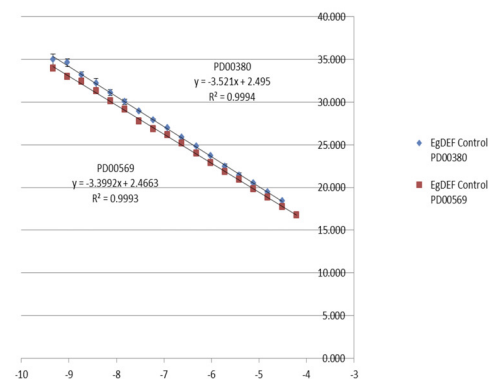
d

Primer	Sequence
1	CTTTCGACACCAAGACCA
2	GTTTCGTCACAGAAAATACCATGT
3	CAAGTAGCGGATAGAGAGGCTTAC
4	TCTTCTGATCGCCTTGCTAGA
5	ACCGGATCAAAGACCGTAAAG

e

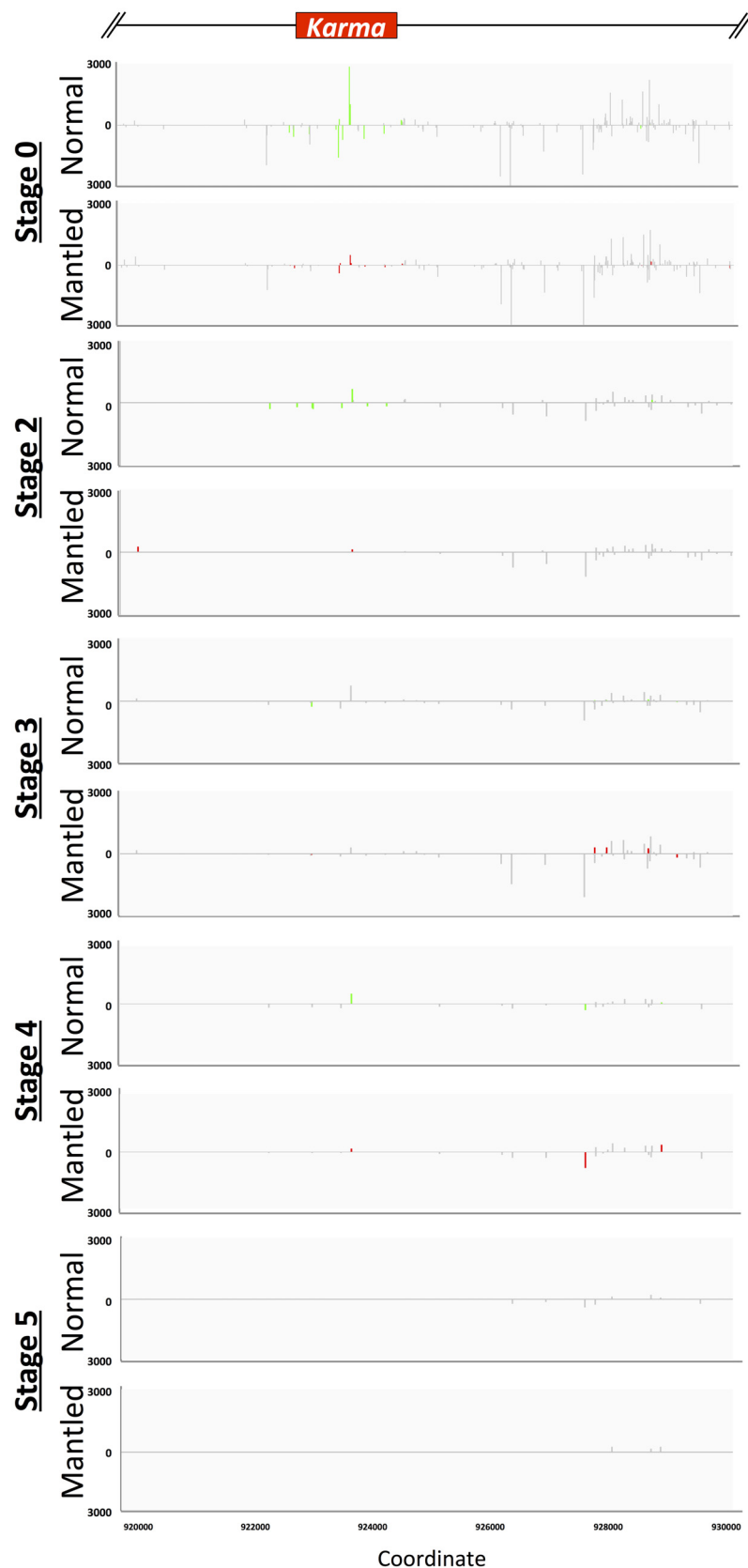


f



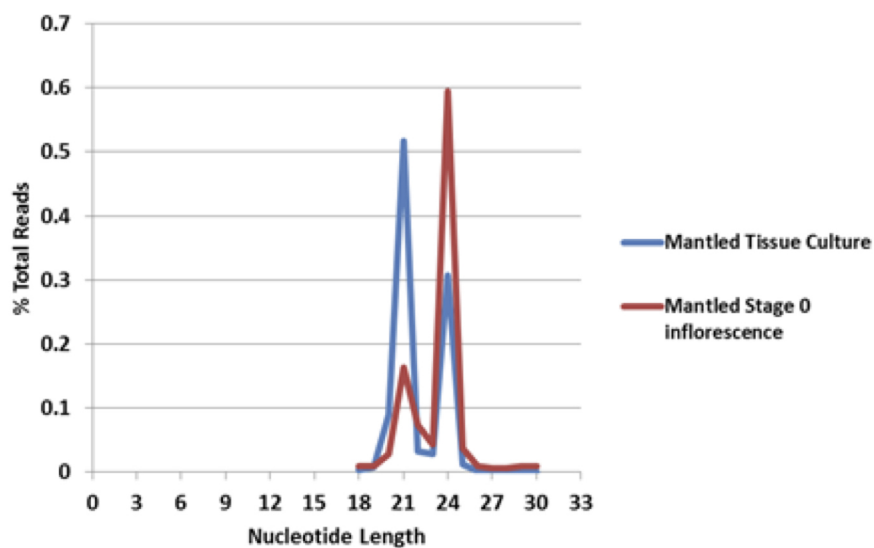
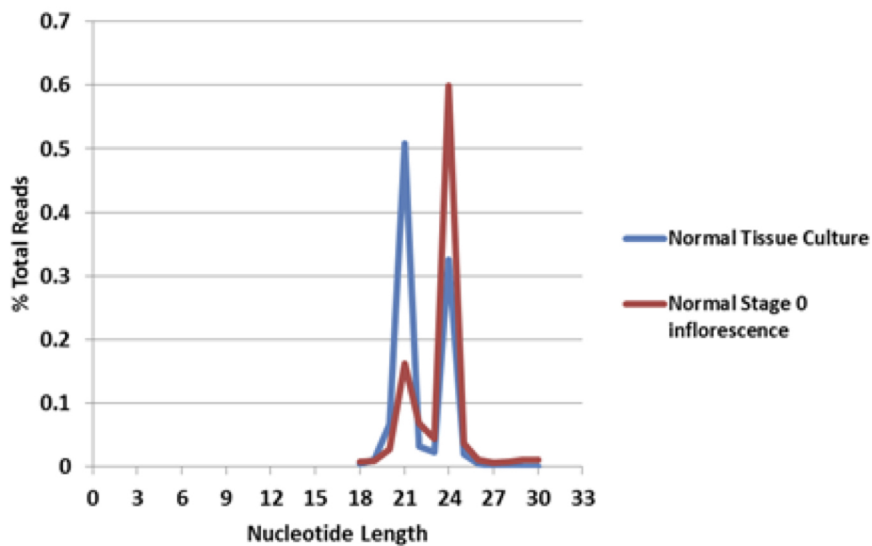
Extended Data Figure 7 | Protein sequences and summary of qRT-PCR assay designs. **a**, Residues highlighted in red are encoded by *Karma* sequence splice to exon 5 of *EgDEF1*. The alternate splicing event disrupts the transcription activation domain of *EgDEF1*. Twelve variant amino acids are coded by *Karma* sequencing, followed by a stop codon. **b**, Diagram of *EgDEF1* locus including positions of qRT-PCR primers. *cDEF1* transcripts were detected using primer a (spanning the splice junction of exons 1 and 2) and primer c (internal to exon 7). *kDEF1* transcripts were detected using primer b (spanning the splice junction of exons 4 and 5) and primer d (internal to *Karma* ORF2). *tDEF1* transcripts were detected using primer a and primer e (spanning the 3' end of exon 5 and including *tDEF1*-specific intron 5 sequence). **c**, All assays were confirmed to give a single band of the correct size by agarose gel electrophoresis. Amplicons were Sanger sequence verified. Note that no band is amplified using the *kDEF1* primer pair in samples from normal inflorescence,

consistent with lack of expression of *kDEF1* in normal inflorescence. **d**, Sequences of primers diagrammed in **b**. **e**, **f**, Standard curves for qRT-PCR assays. PCR amplicons including each qRT-PCR amplified sequence were serially diluted and quantified in triplicate by qPCR using the indicated primer pairs. Dilutions (*x* axis) were plotted against the measured cycle threshold (*y* axis). **e**, Standard curves for *cDEF1* (blue), *kDEF1* (red) and *tDEF1* (green). Line equations were used to calculate the efficiency of each primer pair. The efficiency of each primer pair was used in calculations for quantification of expression of each associated transcript. **f**, Standard curves for two endogenous oil palm control genes. The efficiency of each primer pair was used in calculations for quantification of expression of each associated transcript. Expression of each alternative transcript was calculated relative to the control PD00569 control. Control qRT-PCR primers are described previously³⁹.



Extended Data Figure 8 | Antisense 24-nucleotide siRNA analysis of inflorescence development. siRNA expression at inflorescence stages 0 (shoot apical meristem), 2, 3, 4 and 5, was analysed by Illumina siRNA sequencing (Methods). FPKM normalized expression values for each measured 24-nucleotide siRNA are plotted in scale with the genomic elements diagrammed at the top of the figure. Grey bars indicate detected 24-nucleotide

siRNAs that are not significantly differentially expressed between normal relative to mantled tissues ($P > 0.05$, Student's t -test, two-tailed assuming equal variance). Differentially expressed 24-nucleotide siRNAs are plotted as green or red bars for normal or mantled tissues, respectively. Bars above and below the zero line represent sense and antisense siRNAs, respectively, and are plotted on the same scale in both directions.

a**b**

Extended Data Figure 9 | Relative abundance of 21- and 24-nucleotide sRNA in normal and mantled reclones and stage 0 inflorescence.

a. Distribution of sRNA lengths derived from mantled reclone (blue) and stage 0 inflorescence (red). **b.** Distribution of sRNA lengths derived from

normal reclone (blue) and stage 0 inflorescence (red). Read lengths of sRNA sequencing reads are plotted as the percentage of total reads for each incremental sRNA nucleotide length.

Lineage	Ramet Phenotype	Recloned Culture	BbvI	RsaI
1	Normal	SC2	33%	95%
1	Normal	SC7	18%	87%
2	Normal	SC2	76%	68%
2	Normal	SC7	34%	26%
1	Mantled	SC2	24%	22%
1	Mantled	SC7	4%	5%
2	Mantled	SC2	2%	3%
2	Mantled	SC7	1%	2%

Extended Data Figure 10 | CHG methylation in recloned tissue cultures.

Tissue cultures were reconstituted from normal and mantled ramets from two clonal lineages ('clones of clones'). Methylation at three CHG sites across the *Karma* DMR was quantified by qPCR assays at two (SC2) and seven (SC7) passages in tissue culture. Cultures derived from normal ramets displayed

higher CHG methylation than those derived from mantled ramets. In both normal and mantled reclones, CHG methylation generally decreased with time in culture. At SC2, the time point at which 24-nucleotide siRNAs were measured (Fig. 4c), the culture from normal ramet lineage 1 had lost methylation at the BbvI (the site nearest the *Karma* splice acceptor site).

BET inhibitor resistance emerges from leukaemia stem cells

Chun Yew Fong^{1,2,3}, Omer Gilan^{1,2}, Enid Y. N. Lam¹, Alan F. Rubin^{4,5}, Sarah Ftouni¹, Dean Tyler^{1,2}, Kym Stanley¹, Devbarna Sinha¹, Paul Yeh^{1,2,3}, Jessica Morison⁶, George Giotopoulos⁶, Dave Lugo⁷, Philip Jeffrey⁷, Stanley Chun-Wei Lee⁸, Christopher Carpenter⁹, Richard Gregory⁷, Robert G. Ramsay^{1,2}, Steven W. Lane¹⁰, Omar Abdel-Wahab⁸, Tony Kouzarides¹¹, Ricky W. Johnstone^{1,2}, Sarah-Jane Dawson^{1,2}, Brian J. P. Huntly⁶, Rab K. Prinjha⁷, Anthony T. Papenfuss^{1,2,4,5} & Mark A. Dawson^{1,2,3}

Bromodomain and extra terminal protein (BET) inhibitors are first-in-class targeted therapies that deliver a new therapeutic opportunity by directly targeting bromodomain proteins that bind acetylated chromatin marks^{1,2}. Early clinical trials have shown promise, especially in acute myeloid leukaemia³, and therefore the evaluation of resistance mechanisms is crucial to optimize the clinical efficacy of these drugs. Here we use primary mouse haematopoietic stem and progenitor cells immortalized with the fusion protein MLL-AF9 to generate several single-cell clones that demonstrate resistance, *in vitro* and *in vivo*, to the prototypical BET inhibitor, I-BET. Resistance to I-BET confers cross-resistance to chemically distinct BET inhibitors such as JQ1, as well as resistance to genetic knockdown of BET proteins. Resistance is not mediated through increased drug efflux or metabolism, but is shown to emerge from leukaemia stem cells both *ex vivo* and *in vivo*. Chromatin-bound BRD4 is

globally reduced in resistant cells, whereas the expression of key target genes such as *Myc* remains unaltered, highlighting the existence of alternative mechanisms to regulate transcription. We demonstrate that resistance to BET inhibitors, in human and mouse leukaemia cells, is in part a consequence of increased Wnt/ β -catenin signalling, and negative regulation of this pathway results in restoration of sensitivity to I-BET *in vitro* and *in vivo*. Together, these findings provide new insights into the biology of acute myeloid leukaemia, highlight potential therapeutic limitations of BET inhibitors, and identify strategies that may enhance the clinical utility of these unique targeted therapies.

Our increasing knowledge of cancer genomes and epigenomes not only implicates epigenetic regulators in the initiation and maintenance of cancer, but also highlights an opportunity for therapeutic intervention^{4,5}. One of the most promising epigenetic therapies to have

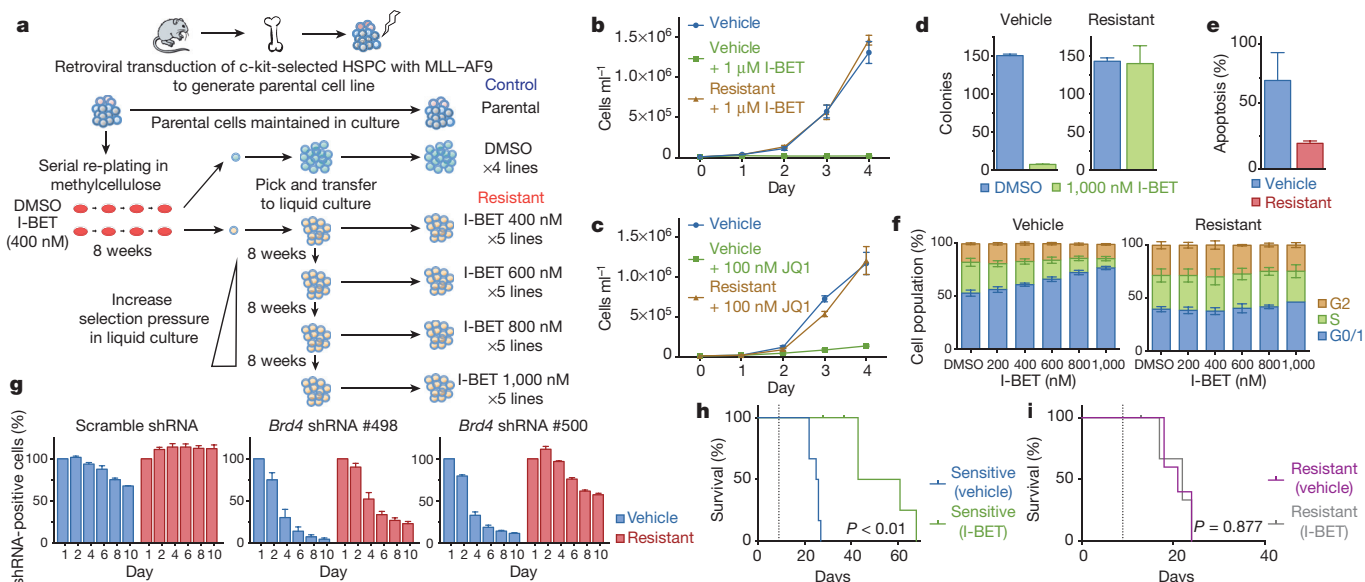


Figure 1 | Establishment of a model of BET inhibitor resistance. **a**, Strategy for the generation of resistant clones. HSPC, haematopoietic stem and progenitor cells. **b**, Resistance to I-BET demonstrated in cell proliferation assays performed in biological triplicate (mean \pm s.d.). **c**, Resistance to JQ1 demonstrated in cell proliferation assays performed in biological triplicate (mean \pm s.d.). **d**, Resistance to I-BET in clonogenic assays performed in biological duplicate (mean \pm s.e.m.). **e**, Resistance to I-BET-mediated

induction of apoptosis in biological triplicate experiments (mean \pm s.e.m.). **f**, Resistant clones do not demonstrate cell cycle arrest in biological triplicate experiments (mean \pm s.e.m.). **g**, Resistance to shRNA-mediated knockdown of Brd4 in biological duplicate experiments (mean \pm s.e.m.). **h**, **i**, Kaplan-Meier curve of secondary syngeneic transplant of sensitive (**h**) and resistant (**i**) clones ($n = 6$ per group, statistical significance calculated using a log-rank test). Dotted line denotes treatment starting on day 9.

¹Cancer Research Division, Peter MacCallum Cancer Centre, East Melbourne, Victoria 3002, Australia. ²Sir Peter MacCallum Department of Oncology, The University of Melbourne, East Melbourne, Victoria 3002, Australia. ³Department of Haematology, Peter MacCallum Cancer Centre, East Melbourne, Victoria 3002, Australia. ⁴Bioinformatics Division, The Walter & Eliza Hall Institute of Medical Research, Parkville, Victoria 3052, Australia. ⁵Department of Medical Biology, University of Melbourne, Parkville, Victoria 3010, Australia. ⁶Department of Haematology, Cambridge Institute for Medical Research and Wellcome Trust-MRC Stem Cell Institute, Cambridge CB2 0XY, UK. ⁷Epinova DPU, Immuno-Inflammation Centre of Excellence for Drug Discovery, GlaxoSmithKline, Medicines Research Centre, Gunnels Wood Road, Stevenage SG1 2NY, UK. ⁸Human Oncology and Pathogenesis Program, Memorial Sloan Kettering Cancer Center, New York, New York 10065, USA. ⁹Cancer Epigenetics DPU, Oncology R&D, GlaxoSmithKline, 1250 South Collegeville Road, Collegeville, Pennsylvania 19426, USA. ¹⁰QIMR Berghofer Medical Research Institute, University of Queensland, Brisbane, Queensland 4029, Australia. ¹¹Gurdon Institute and Department of Pathology, Tennis Court Road, Cambridge CB2 1QN, UK.

emerged in the past decade are small molecule inhibitors targeting the bromodomains of BET family proteins (BRD2, BRD3, BRD4 and BRDT)^{1,2}. While these non-catalytic inhibitors are currently being evaluated in clinical trials across a range of malignancies, the molecular and cellular mechanisms that govern sensitivity and resistance remain largely unknown. We and others have previously demonstrated the pre-clinical efficacy of BET inhibitors in acute myeloid leukaemia (AML)^{6–8}, and early clinical evidence has reinforced the potential of these drugs³.

To study BET inhibitor resistance in a model of AML, we transduced mouse bone marrow haematopoietic stem and progenitor cells (HSPCs) with MLL–AF9. After a selection period in cytokine-supplemented methylcellulose in the presence of dimethylsulfoxide (DMSO; vehicle) or I-BET at the IC₄₀ value (40% of maximal inhibitory effect concentration) of these cells (400 nM), we isolated individual blast colonies, each derived from a single cell, to generate four independent vehicle-treated and five independent I-BET-resistant cell lines (Fig. 1a). The selection pressure on I-BET-resistant clones was sequentially increased to establish clones stably growing at various concentrations including those greater than the IC₉₀ value of the parental and vehicle-treated cells (Fig. 1a, b and Extended Data Fig. 1a). While chemically distinct inhibitors directed against the same target

have sometimes overcome resistance⁹, our data indicates that resistance to I-BET also confers cross-resistance to the chemically distinct BET inhibitor JQ1 (ref. 10) (Fig. 1c and Extended Data Fig. 1b).

Direct comparison of these cell lines demonstrated that although vehicle-treated cells remained exquisitely sensitive to I-BET-mediated suppression of clonogenic capacity, induction of apoptosis and cell cycle arrest, the resistant cells were now impervious to these established phenotypic responses at levels that positively correlated with the degree of selective pressure applied (Fig. 1d–f and Extended Data Fig. 1c). High-content short hairpin RNA (shRNA) screens in this AML model previously identified Brd4 as the major therapeutic target of BET inhibitors⁸. Using an inducible shRNA system, we were able to replicate these findings in our vehicle-treated clones; however, BET-inhibitor-resistant clones were significantly less susceptible to genetic depletion of Brd4 (Fig. 1g and Extended Data Fig. 1d–h).

Consistent with our previous data⁷, I-BET leads to a significant survival advantage in this AML model (Fig. 1h). By contrast, this survival advantage is abrogated following an identical treatment strategy in recipients of resistant cells (Fig. 1i). No differences in morphology or pattern of disease between sensitive or resistant cells were observed (Extended Data Fig. 1i and data not shown). Together, these findings establish a robust model of BET inhibitor resistance *in vitro*

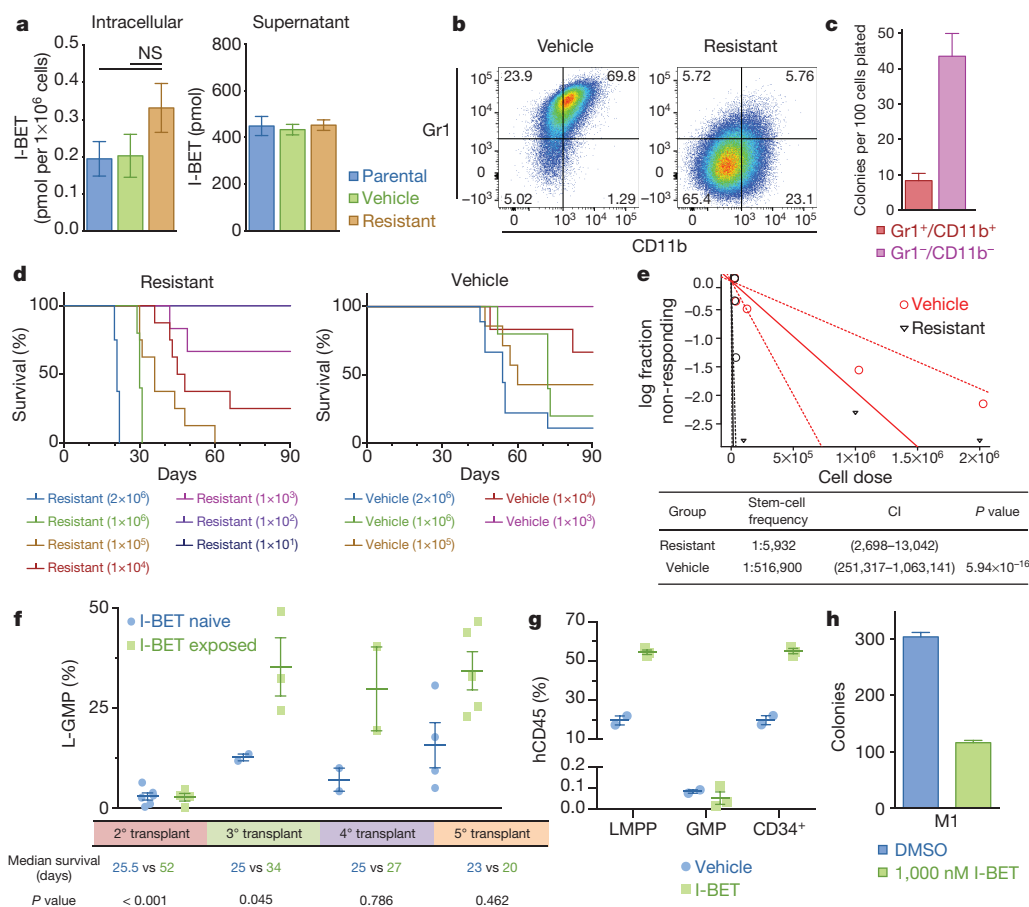


Figure 2 | Resistance to BET inhibitors arises from the LSC compartment.

a, Intracellular and extracellular concentrations of I-BET as assessed by quantitative mass spectrometry in biological duplicates (mean ± s.e.m., statistical significance calculated using a two-tailed Student's *t*-test). NS, not significant. **b**, Resistant clones demonstrate an immature immunophenotype (Gr1⁺/CD11b⁺). **c**, Clonogenic capacity of the Gr1⁺/CD11b⁺ and Gr1⁺/CD11b⁺ populations in resistant clones performed in biological duplicate (mean ± s.e.m.). **d**, Limiting dilution transplantation analyses Kaplan–Meier curves of C57BL/6 mice injected with indicated number of cells; detailed cohort and survival data can be found in Extended Data Fig. 3a. **e**, LSC

frequency from limiting dilution transplantation analyses. Dotted lines indicate 95% confidence intervals (CI). **f**, L-GMP frequency in whole mouse bone marrow (mean ± s.e.m.) after serial transplantation of I-BET-exposed leukaemias from *in vivo* resistance model. Statistical significance of survival outcomes determined using log-rank test of Kaplan–Meier survival estimates. **g**, Proportion of human leukaemic CD34⁺ cells, GMPs and LMPPs in whole mouse bone marrow (mean ± s.e.m.) after I-BET exposure in an AML PDX model (*n* = 5). **h**, I-BET-naïve L-GMPs do not demonstrate intrinsic resistance to I-BET in clonogenic assays in biological duplicate experiments (mean ± s.e.m.), see also Extended Data Fig. 5b. M1, mouse 1.

and *in vivo*, and show that resistant cells are refractory to either chemical or genetic perturbation of Brd4.

Major mechanisms of drug resistance include reduced drug influx or increased drug efflux¹¹. To address this issue, we performed quantitative mass spectrometry, which revealed no significant difference in the amount of intracellular or extracellular drug (Fig. 2a). However, we noted that resistant cells were smaller and more homogenous by flow cytometry (Extended Data Fig. 1j), and further immunophenotypic characterization of sensitive and resistant cells revealed marked differences in the expression of the lineage markers Gr1 and CD11b (Fig. 2b). These findings were replicated in an independent MLL-ENL model of BET inhibitor resistance (Extended Data Fig. 2).

While the precise immunophenotype of leukaemia stem cells (LSCs) in mouse MLL leukaemia models has been debated^{12–14}, it has previously been shown that LSC potential primarily resides in the more immature, lineage-negative (Lin^- , Sca^- , cKit^+ , CD34^+ , $\text{Fc}\gamma\text{R}/\text{RII}^+$) leukaemic granulocyte-macrophage progenitor (L-GMP) population, raising the possibility that BET-inhibitor-resistant cells are enriched for LSCs^{12,14,15}. Consistent with this notion, we noted a significant increase in the blast colony forming potential of the Lin^- ($\text{Gr1}^+/\text{CD11b}^-$) population, and a marked increase in L-GMP cells in our resistant

population before primary transplantation (Fig. 2c and Extended Data Fig. 1k).

While primary transplantation of vehicle-treated cells paralleled the natural history of this AML model, remarkably, primary transplantation of I-BET-resistant cells resulted in considerably shorter leukaemia latency (Fig. 2d). Moreover, limiting dilution transplantation analyses confirm that I-BET-resistant cells were markedly enriched for LSC potential (Fig. 2d, e and Extended Data Fig. 3a). To assess the relevance of these findings to resistance that emerges *in vivo* after sustained exposure to I-BET, we derived an independent *in vivo* model of I-BET resistance (Extended Data Fig. 3b, c). These data validated findings from the *ex vivo* model, and show that *in vivo* BET-inhibitor resistance also emerges from an L-GMP/LSC population (Fig. 2f and Extended Data Fig. 3d–g). Importantly, these I-BET-resistant AML cells have a functional LSC frequency of approximately 1:6; this is virtually identical to what has previously been reported for a purified L-GMP population¹².

To extend these findings into primary patient samples we treated a patient-derived xenograft (PDX) model of AML with I-BET. While the immunophenotype of human AML LSCs can be variable¹⁶, several PDX models have shown that LSCs are enriched within CD34^+

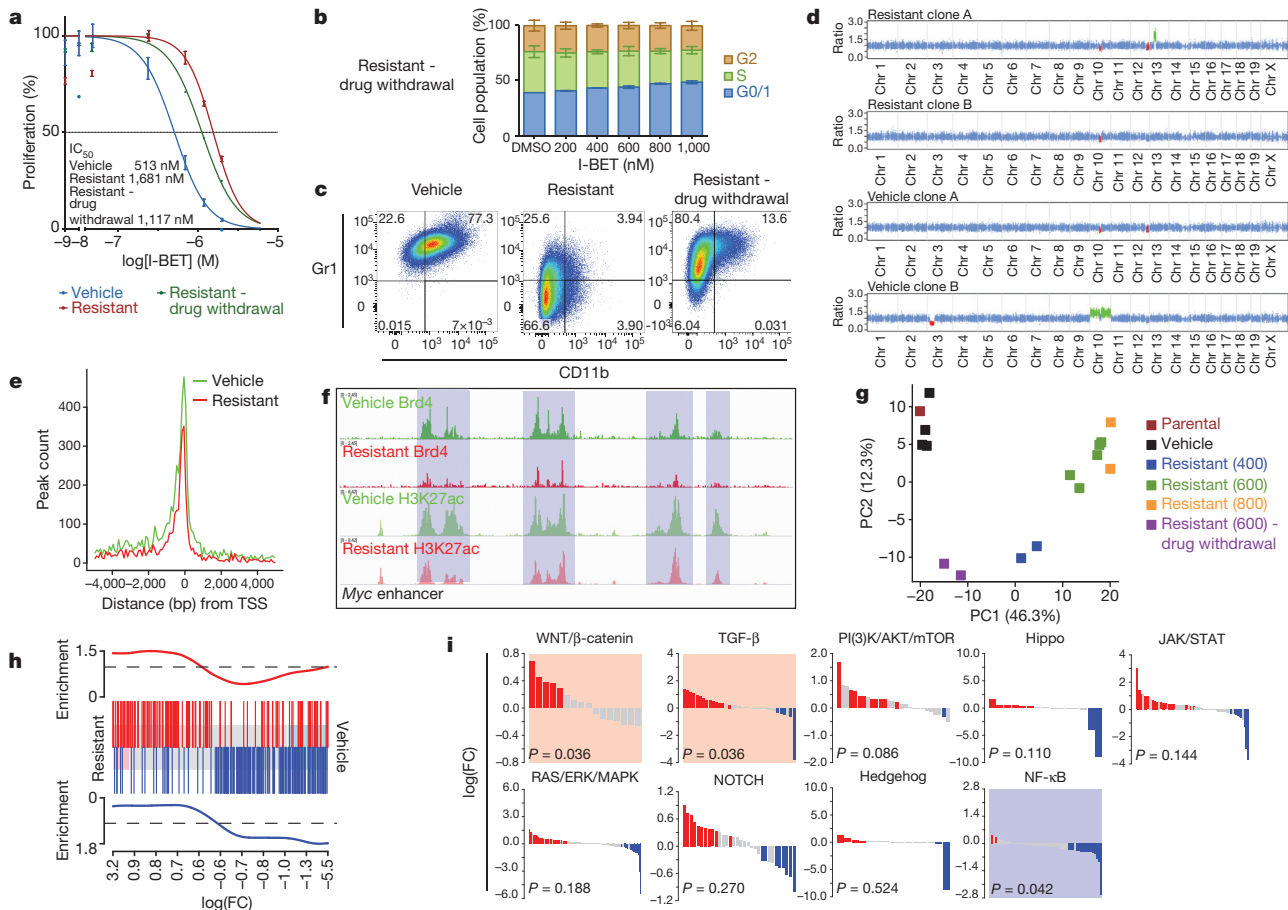


Figure 3 | Genetic, epigenetic and transcriptional characterization of BET-inhibitor-resistant cells. **a**, Proliferation assays in sensitive, resistant cells maintained in I-BET and after 8 weeks drug withdrawal (mean \pm s.d., $n = 12$ per group). **b**, Cell cycle profile in resistant clones after drug withdrawal performed in biological triplicate experiments (mean \pm s.e.m.). **c**, Immunophenotype of sensitive, resistant and drug-withdrawal cells. **d**, Whole-exome capture sequencing data from vehicle-treated and resistant clones normalized to the parental cell line; red regions denote copy number loss and green regions denote copy number gain. **e**, Brd4 binding profile across all annotated transcriptional start sites (TSSs). **f**, Brd4 binding and histone 3 Lys 27 acetylation (H3K27ac) at *Myc* enhancer elements. **g**, Principle component (PC)

analysis of parental cells, vehicle-treated clones ($n = 4$), resistant clones ($n = 9$) and resistant clones after drug withdrawal ($n = 2$). Parentheses denote concentration of I-BET (nM) in which resistant clones have been stably maintained. **h**, GSEA identifies enrichment of a published LSC signature in resistant clones. Upregulated and downregulated genes in the published LSC signature are shown in red and blue, respectively, and correlate with upregulated and downregulated (false discovery rate (FDR) $< 5.0 \times 10^{-5}$) genes in the I-BET-resistant clones. **i**, Statistically significant upregulation (shaded red) of the WNT/ β -catenin and TGF- β pathways and downregulation (shaded blue) of the NF- κ B pathway is observed in all resistant clones ($n = 11$).

cells^{16,17}, which immunophenotypically parallel GMPs or lymphoid-primed multipotent progenitors (LMPPs)¹⁸. Consistent with the data from our mouse AML models, we find that I-BET treatment enriches for the leukaemic LMPP population (Fig. 2g and Extended Data Fig. 4).

To understand whether LSCs were intrinsically resistant to I-BET, we sorted L-GMPs from mice that were I-BET-naïve, and challenged them with 1 μ M of I-BET in clonogenic assays. While this dose virtually eradicates the clonogenic potential of I-BET-naïve bulk leukaemia cells (Fig. 1d and ref. 7), between 30 and 40% of L-GMPs are able to survive (Fig. 2h and Extended Data Fig. 5). Moreover, initial treatment with I-BET *in vivo* does not result in an immediate increase in L-GMPs, instead this population progressively emerges with continuous and sustained exposure to drug *in vivo* (Fig. 2f). These findings suggest that immunophenotypically homogenous L-GMPs/LSCs show marked heterogeneity in their response to I-BET, and that not all L-GMPs are intrinsically resistant to BET inhibitors.

We next sought to understand whether BET inhibitor resistance was reversible in the absence of continuing selective pressure with I-BET. Surprisingly we find that BET inhibitor sensitivity was only partially restored (Fig. 3a, b), and these cells only partially reacquire the immunophenotype of sensitive I-BET-naïve cells (Fig. 3c). Moreover, transcriptionally, they also adopt an intermediate state between sensitive cells and those resistant to I-BET above the IC₆₀ value of the drug (Fig. 3g).

To explore the molecular aetiology for BET inhibitor resistance further, we initially performed whole-exome sequencing in the parental and two separate vehicle/I-BET-resistant cell lines (Fig. 3d and Extended Data Fig. 6). Similar to human leukaemias driven by MLL

fusion proteins¹⁹, these mouse leukaemia cells do not demonstrate significant genomic instability (Extended Data Fig. 6). Notably, although independently established resistant clones behaved identically in all functional analyses described above, there were no gate-keeper mutations in the bromodomains of *Brd2/3/4*, and no shared copy number aberrations. Moreover, only a few mutations with no apparent functional relevance to AML and/or BET activity were shared across several resistant cell lines (Fig. 3d and Extended Data Fig. 6a, b).

We, and others, have shown that treatment with BET inhibitors results in incomplete displacement of Brd2, Brd3 and Brd4 from chromatin^{6,20}. Similarly, we noticed that resistant cells stably growing in I-BET also showed a decrease in chromatin-bound Brd2, Brd3 and Brd4 (Fig. 3e and Extended Data Fig. 7a, b). Notably, however, we found that key Brd4 target genes such as *Myc* were equally expressed in resistant cells despite loss of Brd4 from functional *Myc* enhancer elements (Figs 3f and 4g). These findings raised the prospect that alternative compensatory transcriptional programmes were active in BET-resistant cells.

Global transcriptome analyses using two distinct methodologies showed a very high degree of correlation, and highlighted several transcriptional changes that clearly distinguished sensitive from resistant cells (Fig. 3g and Extended Data Fig. 7c, d). Notably, and consistent with our functional data, gene set enrichment analyses (GSEA) of our resistant cells strongly overlapped with previously published transcriptome data of LSCs from this AML model^{12,15} (Fig. 3h and Extended Data Fig. 7e–j). To identify precise transcriptional

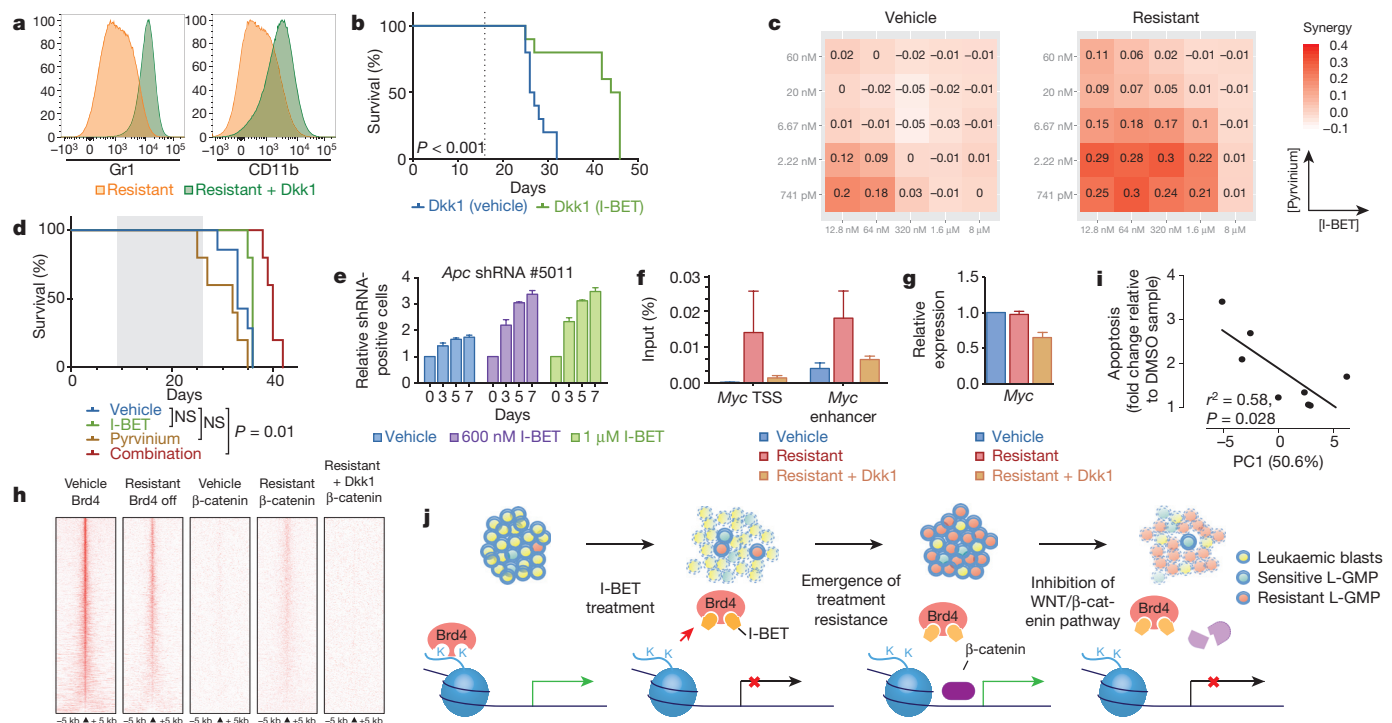


Figure 4 | WNT/β-catenin signalling regulates sensitivity to BET inhibition. **a**, Dkk1 results in re-expression of differentiation markers (Gr1⁺/CD11b⁺). **b**, Kaplan–Meier curve of vehicle and I-BET-treated mice after syngeneic transplantation of resistant clone stably transduced with Dkk1 ($n = 10$ per group, statistical significance calculated using a log-rank test). Dotted line denotes treatment starting on day 16. **c**, Heat map representation of Bliss interaction index across five-point dose range of pyrimin and I-BET performed in biological quadruplicate. **d**, Kaplan–Meier curve of I-BET with/without pyrimin after syngeneic transplantation of resistant cells. Shaded area denotes active treatment between days 9 and 26 (vehicle $n = 7$, I-BET $n = 5$, pyrimin $n = 5$, combination $n = 5$, statistical significance calculated using a log-rank test). **e**, Viable, shRNA-positive cells after treatment

with either vehicle or I-BET normalized to day 0 performed in biological quadruplicate (mean \pm s.d.). **f**, Binding of β-catenin at *Myc* TSS and enhancer elements in vehicle-treated cells, resistant cells and resistant cells with Dkk1. Mean enrichment relative to input (\pm s.e.m.) in chromatin immunoprecipitation (ChIP) analysis from biological triplicate experiments. **g**, *Myc* expression from biological triplicate experiments (mean \pm s.d.). **h**, Heat map representation of Brd4 and β-catenin chromatin occupancy ranked according to amount of Brd4 binding in vehicle-treated clones. **i**, Correlation between aggregate relative expression of examined β-catenin pathway genes with responsiveness to I-BET therapy. Statistical significance determined using Pearson's correlation. **j**, Schematic model proposing the mechanism of resistance to BET inhibitor therapy in AML.

programmes differentially expressed, we performed GSEA for major signalling pathways. These findings demonstrated that the NF- κ B pathway was significantly downregulated, whereas both the TGF- β and Wnt/ β -catenin pathways were significantly upregulated in our resistant cells (Fig. 3i).

We focused our attention on the Wnt/ β -catenin pathway as several components from ligand receptors to transcriptional co-activators were noted to be transcriptionally upregulated (Extended Data Fig. 8a). Interestingly, this pathway has previously been shown to be a major protagonist involved in sustaining LSCs in these models of AML^{14,21} and in other cancer stem cells²² (Extended Data Fig. 8b). To antagonise Wnt/ β -catenin signalling specifically, we overexpressed the Dickkopf Wnt signalling pathway inhibitor 1 (Dkk1), which resulted in the differentiation of our resistant cells into more mature leukaemic blasts (Fig. 4a and Extended Data Fig. 8f) and re-instated sensitivity to I-BET both *in vitro* and *in vivo* (Fig. 4b and Extended Data Fig. 8c–h). In support of these findings, pyrvinium, an established inhibitor of the Wnt/ β -catenin pathway²³, phenocopied these results (Fig. 4c, d and Extended Data Fig. 8i–m). Importantly, stimulation of the Wnt/ β -catenin pathway in sensitive cells, by downregulation of the adenomatous polyposis coli (*Apc*) gene, confers rapid I-BET resistance (Fig. 4e and Extended Data Fig. 9), further highlighting the crucial influence of this pathway on BET inhibitor efficacy.

Mechanistically, we find that in I-BET-naïve cells, Brd4 is bound to the *cis*-regulatory elements of target genes such as *Myc* (Fig. 3f), whereas β -catenin is essentially absent (Fig. 4f). However, in I-BET-resistant cells, Brd4 binding is decreased (Fig. 3f), but β -catenin is now bound at these sites and able to sustain the expression of *Myc* (Fig. 4f, g). Negative regulation with Dkk1 reduces chromatin-bound β -catenin and subverts its ability to maintain the expression of *Myc* (Fig. 4f, g and Extended Data Fig. 8h). Analogous to the events at *Myc*, we find that in the resistant cells, chromatin occupancy of β -catenin increases at the sites where Brd4 is displaced from chromatin, and this increased β -catenin occupancy is abrogated by the expression of Dkk1 (Fig. 4h).

We have previously shown that BET inhibitors have a broad range of efficacy against human AML samples^{6,7}. To explore the translational relevance of our findings, we compared baseline expression of WNT/ β -catenin target genes to the degree of I-BET-induced apoptosis in these samples (Extended Data Fig. 10a). Notably, we find a high degree of correlation (Fig. 4i and Extended Data Fig. 10), supporting our findings that increased activity of the WNT/ β -catenin pathway negates the effects of BET inhibition.

New classes of anti-cancer therapy rarely emerge, and BET inhibitors have uncovered a new therapeutic precedent; the possibility of specifically targeting epigenetic readers (effector proteins that recognize specific epigenetic modifications on histones or nucleotides). If their early clinical promise is to be realized, it is imperative that we evaluate their limitations and mechanisms of resistance to identify rational strategies that enhance their efficacy. Using models that have recapitulated the hierarchical structure of AML *in vitro* and *in vivo*, we show that BET inhibitor resistance emerges from LSCs with increased expression of the Wnt/ β -catenin pathway. While not all LSCs are intrinsically resistant, it is clear that a small proportion of these are either transcriptionally primed or display rapid transcriptional plasticity to survive the initial BET inhibitor challenge, these cells subsequently thrive and become the dominant population (Fig. 4j). This adaptive transcriptional plasticity is an emerging theme by which malignant cells are able to escape from therapeutic pressures²⁴, and our findings are consistent with another report highlighting the WNT/ β -catenin pathway as a mechanism to circumvent BET inhibition²⁵. Our approach has allowed us to sustain a highly enriched population of LSCs in culture indefinitely, providing a unique resource to characterize LSCs molecularly and enable screening of a range of therapies that may ultimately deliver the opportunity to eradicate the LSC population.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 4 November 2014; accepted 3 July 2015.

Published online 14 September 2015.

- Dawson, M. A., Kouzarides, T. & Huntly, B. J. Targeting epigenetic readers in cancer. *N. Engl. J. Med.* **367**, 647–657 (2012).
- Shi, J. & Vakoc, C. R. The mechanisms behind the therapeutic activity of BET bromodomain inhibition. *Mol. Cell* **54**, 728–736 (2014).
- Herait, P. E. *et al.* BET-bromodomain inhibitor OTX015 shows clinically meaningful activity at nontoxic doses: interim results of an ongoing phase I trial in hematologic malignancies. *Cancer Res.* **74**, CT231 (2014).
- Dawson, M. A. & Kouzarides, T. Cancer epigenetics: from mechanism to therapy. *Cell* **150**, 12–27 (2012).
- Helin, K. & Dhanak, D. Chromatin proteins and modifications as drug targets. *Nature* **502**, 480–488 (2013).
- Dawson, M. A. *et al.* Recurrent mutations, including NPM1c, activate a BRD4-dependent core transcriptional program in acute myeloid leukemia. *Leukemia* **28**, 311–320 (2014).
- Dawson, M. A. *et al.* Inhibition of BET recruitment to chromatin as an effective treatment for MLL-fusion leukaemia. *Nature* **478**, 529–533 (2011).
- Zuber, J. *et al.* RNAi screen identifies Brd4 as a therapeutic target in acute myeloid leukemia. *Nature* **478**, 524–528 (2011).
- Weisberg, E., Manley, P. W., Cowan-Jacob, S. W., Hochhaus, A. & Griffin, J. D. Second generation inhibitors of BCR-ABL for the treatment of imatinib-resistant chronic myeloid leukaemia. *Nature Rev. Cancer* **7**, 345–356 (2007).
- Filippakopoulos, P. *et al.* Selective inhibition of BET bromodomains. *Nature* **468**, 1067–1073 (2010).
- Holohan, C., Van Schaeybroeck, S., Longley, D. B. & Johnston, P. G. Cancer drug resistance: an evolving paradigm. *Nature Rev. Cancer* **13**, 714–726 (2013).
- Krivtsov, A. V. *et al.* Transformation from committed progenitor to leukaemia stem cell initiated by MLL-AF9. *Nature* **442**, 818–822 (2006).
- Somervaille, T. C. & Cleary, M. L. Identification and characterization of leukemia stem cells in murine MLL-AF9 acute myeloid leukemia. *Cancer Cell* **10**, 257–268 (2006).
- Wang, Y. *et al.* The Wnt/ β -catenin pathway is required for the development of leukemia stem cells in AML. *Science* **327**, 1650–1653 (2010).
- Krivtsov, A. V. *et al.* Cell of origin determines clinically relevant subtypes of MLL-rearranged AML. *Leukemia* **27**, 852–860 (2013).
- Valent, P. *et al.* Cancer stem cell definitions and terminology: the devil is in the details. *Nature Rev. Cancer* **12**, 767–775 (2012).
- Eppert, K. *et al.* Stem cell gene expression programs influence clinical outcome in human leukemia. *Nature Med.* **17**, 1086–1093 (2011).
- Goardon, N. *et al.* Coexistence of LMP-like and GMP-like leukemia stem cells in acute myeloid leukemia. *Cancer Cell* **19**, 138–152 (2011).
- Andersson, A. K. *et al.* The landscape of somatic mutations in infant MLL-rearranged acute lymphoblastic leukemias. *Nature Genet.* **47**, 330–337 (2015).
- Lovén, J. *et al.* Selective inhibition of tumor oncogenes by disruption of super-enhancers. *Cell* **153**, 320–334 (2013).
- Yeung, J. *et al.* β -Catenin mediates the establishment and drug resistance of MLL leukemic stem cells. *Cancer Cell* **18**, 606–618 (2010).
- Jamieson, C. H. *et al.* Granulocyte-macrophage progenitors as candidate leukemic stem cells in blast-crisis CML. *N. Engl. J. Med.* **351**, 657–667 (2004).
- Thorne, C. A. *et al.* Small-molecule inhibition of Wnt signaling through activation of casein kinase 1 α . *Nature Chem. Biol.* **6**, 829–836 (2010).
- Knoechel, B. *et al.* An epigenetic mechanism of resistance to targeted therapy in T cell acute lymphoblastic leukemia. *Nature Genet.* **46**, 364–370 (2014).
- Rathert, P. *et al.* Transcriptional plasticity promotes primary and acquired resistance to BET inhibition. *Nature* <http://dx.doi.org/10.1038/nature14898> (2015).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank A. Bannister for critical reading of the manuscript. The Leukaemia Foundation Australia, Haematology Society of Australia and New Zealand, Royal Australasian College of Physicians and the Victorian Comprehensive Cancer Centre have supported CYF with PhD scholarships. M.A.D. is a Senior Leukaemia Foundation Australia Fellow, VESKI Innovation Fellow and Herman Clinical Fellow. The National Health and Medical Research Council of Australia (1085015; 1066545) and Leukaemia Foundation Australia fund the Dawson laboratory.

Author Contributions C.Y.F. and M.A.D. designed the research, interpreted data and wrote the manuscript. C.Y.F., O.G., E.Y.N.L., A.F.R., S.F., D.T., K.S., D.S., P.Y., J.M., G.G., D.L., R.G., A.T.P. and M.A.D. performed experiments and/or analysed data. E.L., A.F.R., P.J., R.G.R., S.C.-W.L., C.C., S.W.L., O.A.-W., T.K., R.W.J., S.-J.D., B.J.P.H., R.K.P. and A.T.P. provided critical reagents, interpreted data and aided in manuscript preparation.

Author Information The data discussed in this publication have been deposited in the NCBI Gene Expression Omnibus (GEO) under accession number GSE63683. Reprints and permissions information is available at www.nature.com/reprints. The authors declare competing financial interests: details are available in the online version of the paper. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to M.A.D. (mark.dawson@petermac.org).

METHODS

Generation of immortalized primary mouse HSPC lines and derivation of clonal cell lines. Initial generation of immortalized parental cell lines was achieved through magnetic bead selection (Miltenyi Biotec) of c-kit positive cells, obtained from whole bone marrow of male and female C57BL/6 mice, and subsequent retroviral transduction with either an MSCV-MLL-AF9-IRES-YFP or an MSCV-MLL-ENL construct.

To generate clonal resistant cell lines, the MLL-AF9-bearing parental cell line was serially re-plated in cytokine-supplemented methylcellulose (Methocult M3434, StemCell Technologies) containing either vehicle (0.1% DMSO) or drug (400 nM I-BET151). Individual vehicle-treated or resistant colonies were picked and transferred to liquid culture to generate clonal cell lines. Resistant cell lines were maintained continuously in drug while being incrementally exposed to increasing concentrations of drug (up to 1 μ M I-BET151). Vehicle treated clones were also continuously maintained in 0.1% DMSO and passaged in identical fashion. The parental cell line was continuously maintained with no exposure to vehicle or drug.

Similarly, to generate resistant cell lines, the MLL-ENL-bearing parental cell line was serially re-plated in cytokine-supplemented methylcellulose containing either vehicle (0.1% DMSO) or drug (400 nM I-BET151). Cells growing in each plate were then washed and transferred to liquid culture to generate cell lines. Resistant cell lines were maintained continuously in drug while being incrementally exposed to increasing concentrations of drug (up to 1 μ M I-BET151). Vehicle-treated clones were also continuously maintained in 0.1% DMSO and passaged in identical fashion. The parental cell line was continuously maintained with no exposure to vehicle or drug.

Cell culture. Primary mouse haematopoietic progenitors and derived cell lines were grown in RPMI-1640 supplemented with mouse IL-3 (10 ng ml⁻¹), 20% FCS, penicillin (100 U ml⁻¹), streptomycin (100 μ g ml⁻¹), amphotericin B (250 ng ml⁻¹) and gentamycin (50 μ g ml⁻¹). Cell lines were routinely tested for mycoplasma contamination by PCR. Primary human leukaemia cells were grown in the presence of IL3 (10 ng ml⁻¹), IL6 (10 ng ml⁻¹) and SCF (50 ng ml⁻¹). Cells were incubated at 37°C and 5% CO₂.

Cell proliferation assays. For dose-response assays, serial dilutions of I-BET151, JQ1 or pyrvinium were further diluted in media before addition to 96-well plates seeded with between 5 × 10³ and 1 × 10⁴ cells per well to obtain a 0.1% DMSO final concentration. After 72 h incubation, resazurin was added to each well and plates were further incubated for 3 h. Fluorescence was then read at 560 nm/590 nm on a Cytation 3 Imaging Reader (BioTek). Cell counts were performed using a haemocytometer. Determination of *in vitro* synergy in proliferation assays was undertaken according to the method described previously²⁶.

Clonogenic assays in methylcellulose. Clonogenic potential was assessed through colony growth of derived cell lines plated in cytokine-supplemented methylcellulose (Methocult M3434, StemCell Technologies). Derived vehicle-treated and resistant cell lines were plated in duplicate at a cell dose of 2 × 10² per plate in the presence of vehicle (0.1% DMSO) or drug (1 μ M I-BET151). Gr1⁺/CD11b⁺ and Gr1⁺/CD11b⁺ fractions of resistant cell lines were plated in duplicate following FACS sorting at a cell dose of between 2 × 10² and 2 × 10³ cells per plate. FACS-isolated L-GMP populations from whole mouse bone marrow following primary syngeneic transplant of vehicle-treated clones were plated in duplicate at a cell dose of between 2 × 10² and 2 × 10³ cells per plate in the presence of vehicle (0.1% DMSO) or drug (1 μ M I-BET151). Cells were incubated at 37°C and 5% CO₂ for 7–10 days at which time colonies were counted.

Flow cytometric analyses. Cell apoptosis was assessed using APC conjugated Annexin V (550475, BD Biosciences) and propidium iodide (PI) (P4864, Sigma-Aldrich) staining according to manufacturer's instructions.

For cell cycle analysis, cells were fixed overnight at -20°C in 70% ethanol/PBS. Before flow cytometry analysis, cells were incubated at 37°C for 30 min in PI staining solution (0.02 mg ml⁻¹ PI, 0.05% (v/v) Triton X-100 in PBS, supplemented with DNase-free RNase A (19101, Qiagen)) or incubated at room temperature for 10 min with 4',6-diamidino-2-phenylindole (DAPI) staining solution (1 μ g ml⁻¹ DAPI, 0.05% (v/v) Triton X-100 in PBS).

Immunophenotype assessment for markers of committed differentiation was undertaken through staining with Alexa Fluor 700 anti-Gr1 (108422, BioLegend) and Brilliant Violet 605 anti-CD11b (101237, BioLegend). Assessment of L-GMP populations was undertaken through staining with eFluor 660 anti-CD34 (50-0341-82, eBioscience), biotin lineage antibody cocktail (120-001-547, Miltenyi Biotec), PerCP/Cy5.5 anti-CD16/32 (101324, BioLegend), APC/Cy7 anti-CD117 (105826, BioLegend) and Pacific Blue anti-Ly-6A (122520, BioLegend) followed by secondary staining with V500 streptavidin (561419, BD Biosciences). Assessment of leukaemic LMPP and GMP populations in patient-derived xenografts was undertaken through staining with APC/Cy7 anti-mouse CD45.1 (110716, BioLegend), eFluor 450 anti-mouse Ter119 (48-5921-82, eBioscience), FITC anti-human CD45

(11-9459-42, eBioscience), BV711 anti-human CD38 (563965, BD Biosciences), PE anti-human CD90 (561970, BD Biosciences), PE-Cy5 anti-human CD123 (15-1239-41, eBioscience), PerCP-Cy5.5 anti-human CD45RA (45-0458-42), biotin anti-human CD3 (555338, BD Biosciences), biotin anti-human CD19 (555411, BD Biosciences), PE-Cy7 anti-human CD33 (333946, BD Biosciences) and APC anti-human CD34 (555824, BD Biosciences) followed by secondary staining with V500 streptavidin (561419, BD Biosciences).

PI or DAPI was used as a viability dye to ensure that immunophenotyping analyses were performed on viable cells. Appropriate unstained, single-stained and fluorescence minus one controls were used to determine background staining and compensation in each channel.

Flow cytometry analyses were performed on a LSRFortessa X-20 flow cytometer (BD Biosciences) and all data analysed with FlowJo software (vX.0.7, Tree Star). Cell sorting was performed on a FACSARIA Fusion flow sorter (BD Biosciences).

RNA interference studies. shRNAs were cloned into TtrMVPVIR (27995, addgene). For competitive proliferation assays, transduced cells were sorted for shRNA-containing (Venus⁺/YFP⁺) and non-shRNA-containing (YFP⁺ only) populations and recombined at a 1:1 ratio. After this, cells were cultured with 1 mg ml⁻¹ doxycycline to induce shRNA expression. The proportion of shRNA-expressing (dsRED⁺/Venus⁺/YFP⁺) cells were determined by flow cytometric analysis and followed over time. Knockdown efficiency of shRNA-expressing and non-shRNA-containing cells was assessed after 48–72 h of doxycycline exposure by quantitative reverse transcriptase PCR (qRT-PCR) and immunoblotting.

The following shRNA sequences were used: *Brd2* (#851), 5'-CGGATTATCA CAAATTAT-3'; *Brd4* (#498), 5'-ACTATGTTTACAAATTGTT-3'; *Brd3/4* (#499), 5'-AGGACTTCAACACTATGTT-3'; *Brd4* (#500), 5'-AGCAGAACAA ACCAAAGAA-3'.

shRNAs directed against *Apc* were cloned into LMN-mirE-mCherry. The proportion of shRNA-expressing (mCherry⁺) cells was determined by flow cytometric analysis following treatment with vehicle (0.1% DMSO) or I-BET151 and followed over time. Selective advantage consequent to shRNA expression results in enrichment of mCherry⁺ cells. Knockdown efficiency of *Apc* in shRNA-expressing cells was assessed following FACS of mCherry⁺ cells. shRNAs directed against *Apc* were a gift from J. Zuber, the detailed validation of which can be found in ref. 25.

qRT-PCR. mRNA was prepared using the Qiagen RNeasy kit and cDNA synthesis was performed using SuperScript VILO kit (Life Technologies) as per manufacturer's instructions. qPCR analysis was undertaken on an Applied Biosystems StepOnePlus System with SYBR green reagents (Life Technologies).

For analysis of mouse cell line samples, expression levels were determined using the ΔC_T method and normalized to β -2-microglobulin (*B2m*) and/or *Gapdh*. Differences in expression were assessed using a one-sided *t*-test for statistical significance. Assessment of expression changes associated with I-BET151 treatment occurred at 6 h after treatment with 1 μ M I-BET151.

The following mouse primer pairs were used: *Apc*, forward 5'-GGAGTGGC AGAAAGCAACAC-3', reverse 5'-AAACACTGGCTGTTCGTGA-3'; *B2m*, forward 5'-GAGCCCAAGACCGTCTACTG-3', reverse 5'-GCTATTTCTTTC TGCGTGCAT-3'; *Brd2*, forward 5'-TGGGCTGCCTCAGATGTAT-3', reverse 5'-CCAGTGTCTGTGCCATTAGG-3'; *Brd3*, forward 5'-GCCAGTGAGTGTA TGCAGGA-3', reverse 5'-GCCTGGGCCATTAGCACTAT-3'; *Brd4*, forward 5'-TCTGCACGACTACTGTGACA-3', reverse 5'-GGCATCTCTGTACTCTC GGG-3'; *Ccnd2*, forward 5'-CAAGCCACCACCCCTACA-3', reverse 5'-TTGC CGCCGAATGG-3'; *Dkk1*, forward 5'-CTGCATGAGGCACGCTATGT-3', reverse 5'-AGGAAAATGGCTGTGGTCAG-3'; *Dvll1*, forward 5'-ATCACAC GCACCAGCTCTTC-3', reverse 5'-GGACAATGGCACTCATGTCA-3'; *Fzd5*, forward 5'-GGCTACAACCTGACGCACAT-3', reverse 5'-CAGAATTGGTG CACCTCCAG-3'; *Gapdh*, forward 5'-GGTGCTGAGTATGTCGTGGA-3', reverse 5'-CGGAGATGATGACCCCTTTG-3'; *Gsk3b*, forward 5'-TTGGAGC CACTGATTACACG-3', reverse 5'-CCAAGTATCCACACCACTG-3'; *Myc*, forward 5'-TGAGCCCCTAGTGCTGCAT-3', reverse 5'-AGCCCGACTCCGA CCTCTT-3'.

For determination of baseline WNT/ β -catenin pathway and target gene expression in primary human AML samples, expression relative to the mean of all samples was determined using the ΔC_T method and normalized to *GAPDH* and actin. The following human primers were used: *AXIN2*, forward 5'-CGGA CAGCAGTGTAGATGGA-3', reverse 5'-CTTCACACTGCGATGCATTT-3'; *CCND1*, forward 5'-GCTGTGCTACACCGACA-3', reverse 5'-CCACTT GAGCTTGTTCACCA-3'; *CTNNB1*, forward 5'-GACCACAAGCAGAGTGC TGA-3', reverse 5'-CTTGCACTCCACCACTCTCT-3'; *FZD5*, forward 5'-TTC CTGTCAGCCTGCTACCT-3', reverse 5'-CGTAGTGGATGTGGTTGTGC-3'; *MYC*, forward 5'-CTGGTGCTCCATGAGGAGA-3', reverse 5'-CCTGCCCTC TTTTCCACAGAA-3'; *TCF4*, 5'-ATGGCAAATAGAGGAAGCGG-3', reverse 5'-TGGAGAATAGATCGAAGCAAG-3'; *ACTB*, forward 5'-TTCAACACCC

CAGCCATGT-3', reverse 5'-GCCAGTGGTACGGCCAGA-3'; *GAPDH*, 5'-ACGGGAAGCTTGTCATCAAT-3', reverse 5'-TGGACTCCACGACGTACTCA-3'.

Immunoblotting. Whole-cell lysates were mixed with Laemmli SDS sample buffer, separated via SDS-PAGE and transferred to PVDF membranes (Millipore). Membranes were then sequentially incubated with primary antibodies (see antibodies) and secondary antibodies conjugated with horseradish peroxidase (Invitrogen). Membranes were then incubated with ECL (GE Healthcare) and proteins detected by exposure to X-ray film.

Mouse tissue sample preparation. Peripheral blood samples were collected in EDTA-treated tubes (Sarstedt) and counted using a XP-100 analyser (Sysmex). Single-cell cytopins and blood smears were stained with the Rapid Romanowsky Staining Kit (Thermo Fisher Scientific). Bone marrow cells were isolated by flushing both femurs and tibias with cold PBS. Before flow cytometry, red blood cells were lysed in red blood cell lysis buffer (Sigma).

Examination of drug efflux and metabolism by quantitative mass spectrometry. Between 2×10^5 and 3×10^5 cells per well were seeded in 24-well plates and treated with vehicle (0.1% DMSO) or 600 nM I-BET151. After 48 h, cells were collected by centrifugation, washed twice in ice-cold PBS and lysed in M-PER buffer (78501, Thermo Scientific). Base media, supernatant, wash and cell lysates were quenched with 5% acetonitrile (aq) containing labetalol at 62.5 ng ml^{-1} as the internal standard. These samples, in addition to serial dilutions of I-BET151 used to generate standard curves, were then analysed by mass spectrometry.

HPLC-mass spectrometry apparatus and conditions: The HPLC system was an integrated CTC PAL auto sampler (LEAP technologies), Jasco XTC pumps (Jasco). The HPLC analytical column was an ACE 2 C18 30 mm \times 2.1 mm (Advanced Chromatography Technologies) maintained at 40 °C. The mobile phase solvents were water containing 0.1% formic acid and acetonitrile containing 0.1% formic acid. A gradient ran from 5% to 95% ACN plus 0.1% formic acid up to 1.3 min, held for 0.1 min and returning to the starting conditions over 0.05 min then held to 1.5 min at a flow rate of 1 ml min^{-1} . A divert valve was used so the first 0.4 min and final 0.2 min of flow were diverted to waste.

Mass spectrometric detection was by an API 4000 triple quadrupole instrument (AB Sciex) using multiple reaction monitoring (MRM). Ions were generated in positive ionization mode using an electrospray interface. The ionspray voltage was set at 4,000 V and the source temperature was set at 600 °C. For collision dissociation, nitrogen was used as the collision gas. The MRM of the mass transitions for I-BET151 (m/z 416.17 to 311.10), and labetalol (m/z 329.19 to 162.00), were used for data acquisition.

Data were collected and analysed using Analyst 1.4.2 (AB Sciex), for quantification, area ratios (between analyte/internal standard) were used to construct a standard line, using weighted ($1/x^2$) linear least squared regression, and results extrapolated the area ratio of samples from this standard line.

Mouse models of leukaemia. Primary syngeneic transplantation studies of stably growing derived vehicle treated or resistant cell lines in limit dilution analyses were performed with intravenous injection of between 1×10^1 to 2×10^6 cells per mouse.

Serial syngeneic transplantation studies of drug efficacy, generation of *in vivo* resistance and limit dilution analyses were performed with intravenous injection of between 1×10^1 to 2.5×10^6 cells per mouse obtained from bone marrow or spleen. Treatment with vehicle or I-BET151 at 20–30 mg kg⁻¹ began between days 9 and 13. Pyrimin, alone or in combination with I-BET151, was delivered between days 9 and 26.

After stable retroviral transduction of resistant cell lines with a Dkk1 containing construct, 5×10^6 cells per mouse were injected intravenously in primary syngeneic transplants. Treatment with vehicle or I-BET151 at 20 mg kg⁻¹ began at day 16.

Syngeneic transplantation studies were performed in C57BL/6 mice (wild-type or expressing Ptprc^o). All mice were 6–10 weeks old at the time of sub-lethal irradiation (300 cGy) and intravenous cell injection. Treatment with vehicle, I-BET151 or pyrimin commenced after engraftment of leukaemia as determined by >1% yellow fluorescent protein (YFP) expression in peripheral blood in most mice. Mice were randomly assigned treatment groups; treatment administration was not blinded. Sample sizes were determined according to the resource equation method. Differences in Kaplan–Meier survival curves were analysed using the log-rank statistic.

Patient derived xenograft studies were performed in NOD/SCID/Il2rg^{-/-} (NSG) mice. All mice were 6–10 weeks old at the time of sub-lethal irradiation (200 cGy) and intravenous cell injection of 1×10^5 to 5×10^5 cells per mouse. Treatment with vehicle or I-BET151 at 10 mg kg⁻¹ for a 2-week period began after detection of >1% circulating human CD45⁺ cells in mouse peripheral blood at week 14. Treatment cohorts were matched for transplant generation.

I-BET151 was dissolved in normal saline containing 5% (v/v) DMSO and 10% (w/v) Kletose HPB. I-BET151 was delivered daily (5 days on, 2 days off) by

intraperitoneal injection (10 ml kg⁻¹) with dose reduction of I-BET151 undertaken if evidence of drug intolerance was present. Pyrimin was dissolved in normal saline containing 15% (v/v) DMSO and delivered daily by intraperitoneal injection (10 ml kg⁻¹). Dosing of pyrimin started at 0.1 mg kg⁻¹ and escalated in 0.1 mg kg⁻¹ increments every second dose to a maximal dose of 0.5 mg kg⁻¹.

All mice were kept in a pathogen-free animal facility, inspected daily and euthanized on signs of distress/disease. All experiments were conducted under either UK Home Office regulations or Institutional Animal Ethics Review Board in Australia. Statistical analyses of limit dilutions were undertaken according to the method described previously²⁷.

Exome capture sequencing. DNA was extracted from cell lines using the DNeasy blood and tissue kit (Qiagen), and quantified using the Qubit dsDNA HS Assay (Life Technologies) before fragmentation to a peak size of approximately 200 base pairs (bp) using the focal acoustic device, SonoLab S2 (Covaris). Library preparations were performed using the SureSelect^{XT} Target Enrichment System for Illumina Paired-End Sequencing Library protocol (Agilent Technologies) with the SureSelect^{XT} Mouse All Exon Kit for the capture process (Agilent Technologies). The quality of libraries submitted for sequencing was assessed using the High Sensitivity DNA assay on the 2100 bioanalyzer (Agilent Technologies). Libraries were quantified with qPCR, normalized and pooled to 2 nM before sequencing with paired end 100-bp reads using standard protocols on the HiSeq2500 (Illumina).

The Fastq files generated by sequencing were aligned to the mm10 mouse reference genome using bwa²⁸. Copy number variation was analysed using ADTEX²⁹ to compare the depth of coverage in resistant and vehicle treated clones with the parental cell line. Variant calling was performed with VarScan2 (ref. 30), Mutect³¹ and GATK HaplotypeCaller³². The Ensembl Variant Effect Predictor (VEP)³³ was used to predict the functional effect of the identified variants.

Mutations detected by at least two variant callers were further analysed for shared mutations between cell lines and mutation spectrum. Genomic regions with coverage of at least eight reads in all libraries were analysed for the frequency of mutations. Coding exonic, untranslated regions and intronic regions were obtained from the UCSC Table Browser³⁴. Upstream regions were defined as 1,000 bp upstream of genes, downstream regions were defined as 1,000 bp downstream of genes, and intergenic regions were more than 1,000 bp from genes.

ChIP, qPCR and sequencing analysis. Cells were cross-linked with 1% formaldehyde for 15 min at room temperature and cross-linking stopped by the addition of 0.125 M glycine. Cells were then lysed in 1% SDS, 10 mM EDTA, 50 mM Tris-HCl, pH 8.0, and protease inhibitors. Lysates were sonicated in a Covaris ultrasonicator to achieve a mean DNA fragment size of 500 bp. Immunoprecipitation (see antibodies) was performed for a minimum of 12 h at 4 °C in modified RIPA buffer (1% Triton X-100, 0.1% deoxycholate, 90 mM NaCl, 10 mM Tris-HCl, pH 8.0 and protease inhibitors). An equal volume of protein A and G magnetic beads (Life Technologies) were used to bind the antibody and associated chromatin. Reverse crosslinking of DNA was followed by DNA purification using QIAquick PCR purification kits (Qiagen). Immunoprecipitated DNA was analysed on an Applied Biosystems StepOnePlus System with SYBR green reagents. The following primer pairs were used in the analysis: *Myc* TSS, forward 5'-GTCACCTTTACCCCGACTCA-3', reverse 5'-TCCAGGCACATCTCAGTTTG-3'; *Myc* enhancer, forward 5'-TCTTTGATGGGCTCAATGGT-3', reverse 5'-TTCCTTCACCTGATGAACC-3'. For sequencing analysis of immunoprecipitated DNA, DNA was quantified using the Qubit dsDNA HS Assay (Life Technologies). Library preparations were performed using the standard ThruPLEXTM-FD Prep Kit protocol (Rubicon Genomics) and size selected for 200–400 bp using the Pippin Prep (Sage Science Inc.). Fragment sizes were established using either the High Sensitivity DNA assay or the DNA 1000 kit and 2100 bioanalyzer (Agilent Technologies). Libraries were quantified with qPCR, normalized and pooled to 2 nM before sequencing with single-end 50-bp reads using standard protocols on the HiSeq2500 (Illumina). The Fastq files generated by sequencing were aligned to the mm10 mouse reference genome using bwa²⁸. Peak-calling was performed using MACS2 (ref. 35) with default parameters and the input library as control. Profiles and heat maps of reads and MACS peaks in the 5 kb around the TSS were generated with Genomic Tools³⁶.

Expression analysis by microarray and RNA-sequencing. RNA was prepared using the Qiagen RNeasy kit. For microarray analysis, RNA was hybridized to Illumina MouseWG-6 v2 Expression BeadChips. Gene expression data were processed using the lumi package in R. Probe sets were filtered to remove those where the detection *P* value (representing the probability that the expression is above the background of the negative control probe) was greater than 0.05 in at least one sample. Expression data was background corrected and quantile normalized. Normalization and inference of differential expression were performed using limma³⁷. Correction for multiple testing was performed using the method of Benjamini and Hochberg³⁸. Genes with an FDR rate below 0.05 and a fold-change greater than 2 were considered significantly differentially expressed. For genes

with multiple probe sets, only the probe set with the highest average expression across samples was used.

For RNA sequencing analysis, RNA concentration was quantified with the NanoDrop spectrophotometer (Thermo Scientific). The integrity was established using the RNA 6000 kit and 2100 bioanalyzer (Agilent Technologies). Library preparations were performed using the standard TruSeq RNA Sample Preparation protocol (Illumina) with fragment sizes established using the DNA 1000 kit and 2100 bioanalyzer (Agilent Technologies). Libraries were quantified with qPCR, normalized and pooled to 2 nM before sequencing with paired-end 50 bp reads using standard protocols on an Illumina HiSeq2500.

Reads were aligned to the mouse genome (Ensembl Release 75, Feb 2014) using Subread³⁹ and assigned to genes using featureCounts⁴⁰. Differential expression was inferred using limma/voom³⁷. Correction for multiple testing using the Benjamini–Hochberg method was performed. Genes with an FDR below 0.05 and a fold-change greater than 2 were considered significantly differentially expressed.

Gene set enrichments were determined using ROAST⁴¹. ROAST tests for up- or downregulation of genes in a given pathway (Fig. 3i) were performed on cell lines either stably maintained in vehicle or I-BET. *P* values were corrected for multiple testing using the method of Benjamini and Hochberg. Gene sets were obtained from MSigDB⁴² and curated. Human Entrez accessions from the downloaded gene sets were converted into mouse accessions using orthologue information from the Mouse Genome Database (MGD) at the Mouse Genome Informatics website (<http://www.informatics.jax.org>; accessed June 2014). ROAST tests were performed to assess for an enrichment of a LGMP gene expression signature (GSE4416)¹² and a LGMP derived from HSC signature (GSE18483)¹⁵ in the I-BET resistant compared with vehicle cell lines. The gene expression program associated with human leukaemia stem cells was obtained from GSE30375 (ref. 17) and analysed with LIMMA³⁷. Gene expression of LSC was compared with LPC and genes upregulated in LSC were analysed for an enrichment of the Wnt/ β -catenin pathway using ROAST.

GSEA terms. The following GSEA terms were used. WNT/ β -catenin: ST_WNT_BETA_CATENIN_PATHWAY; JAK/STAT: KEGG_JAK_STAT_SIGNALING_PATHWAY; PI(3)K/AKT/mTOR: REACTOME_PI3K_AKT_ACTIVATION; NF- κ B: REACTOME_ACTIVATION_OF_NF_KAPPAB_IN_B_CELLS; RAS/ERK/MAPK: KEGG_MAPK_SIGNALING_PATHWAY; NOTCH: KEGG_NOTCH_SIGNALING_PATHWAY; hippo: REACTOME_SIGNALING_BY_HIPPO; hedgehog: KEGG_HEDGEHOG_SIGNALING_PATHWAY; TGF- β : KEGG_TGF_BETA_SIGNALING_PATHWAY.

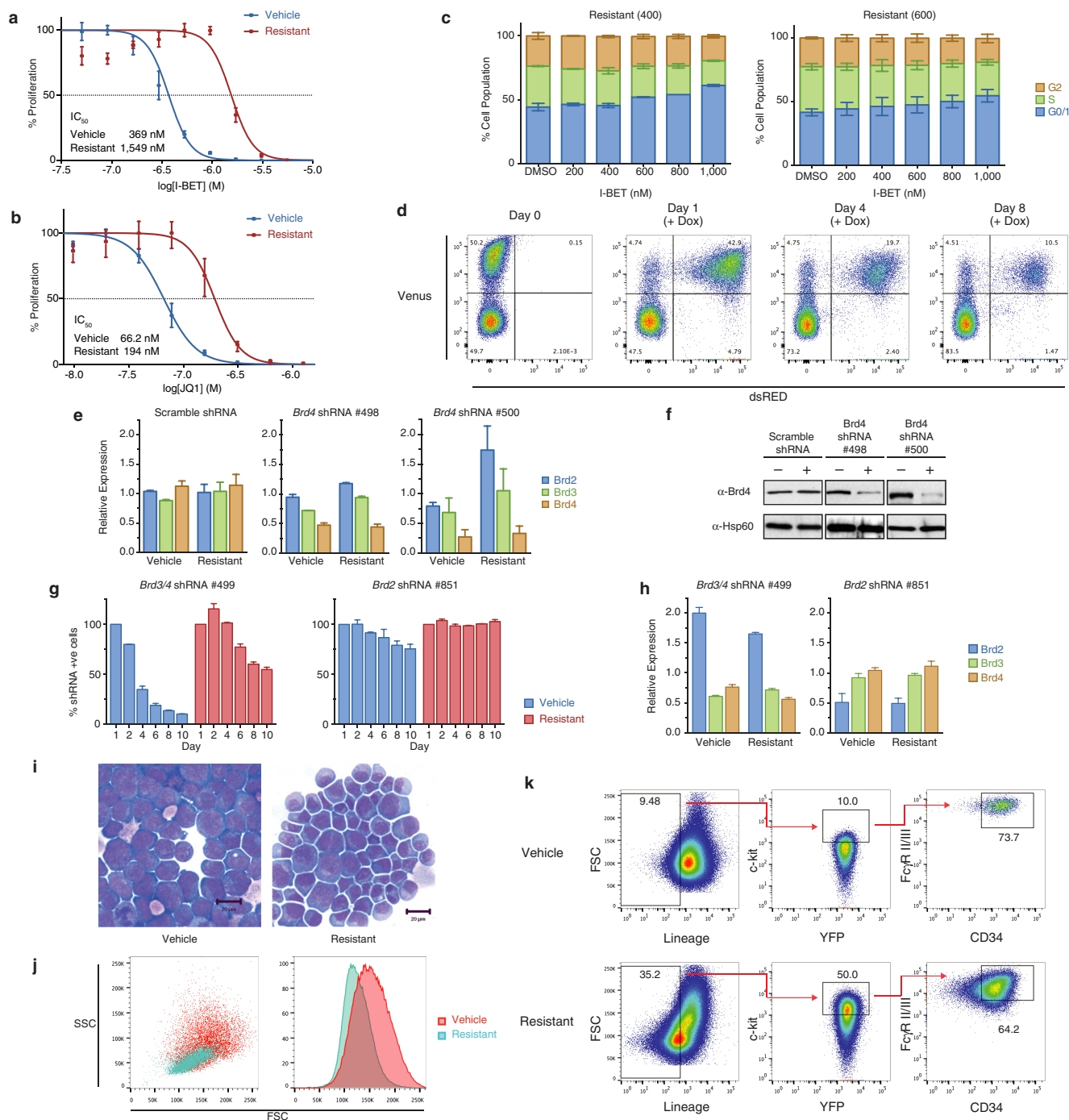
Antibodies. The following antibodies were used in ChIP and immunoblotting assays: anti-Brd2 (A302-583A, Bethyl Labs), anti-Brd3 (A302-368A, Bethyl Labs), anti-Brd4 (A301-985A, Bethyl Labs and ab128874, abcam), anti-H3K27ac (ab4729, abcam), anti- β -catenin (610154, BD Biosciences), anti-c-Myc (9402S, Cell Signalling Technology), anti- β -actin (A1978, Sigma-Aldrich) and anti-Hsp60 (sc-13966, Santa Cruz Biotechnology).

Correlation of expression of WNT/ β -catenin pathway expression and response to I-BET151. A principal component analysis was performed on the qRT-PCR data of β -catenin pathway and target genes from primary human AML samples. Pearson's correlation was calculated between the expression of the pathway genes in the first principal component, and the responsiveness to I-BET151.

Correlation of log gene expression of selected WNT/ β -catenin pathway genes was assessed using a corrgram and correlation between log expression and apoptosis was examined using scatterplots. As expression between genes was typically highly correlated, or inversely correlated, the log-expression data was summarized using the first principle component and compared to the level of apoptosis. A multiple linear regression model was also fitted to the data. As the full model was close to saturated (8 samples, 6 genes), a stepwise model selection procedure based on the Akaike Information Criteria (AIC), which was implemented in the R function STEP, was used. The model that minimized the AIC excluded one gene (AXIN2).

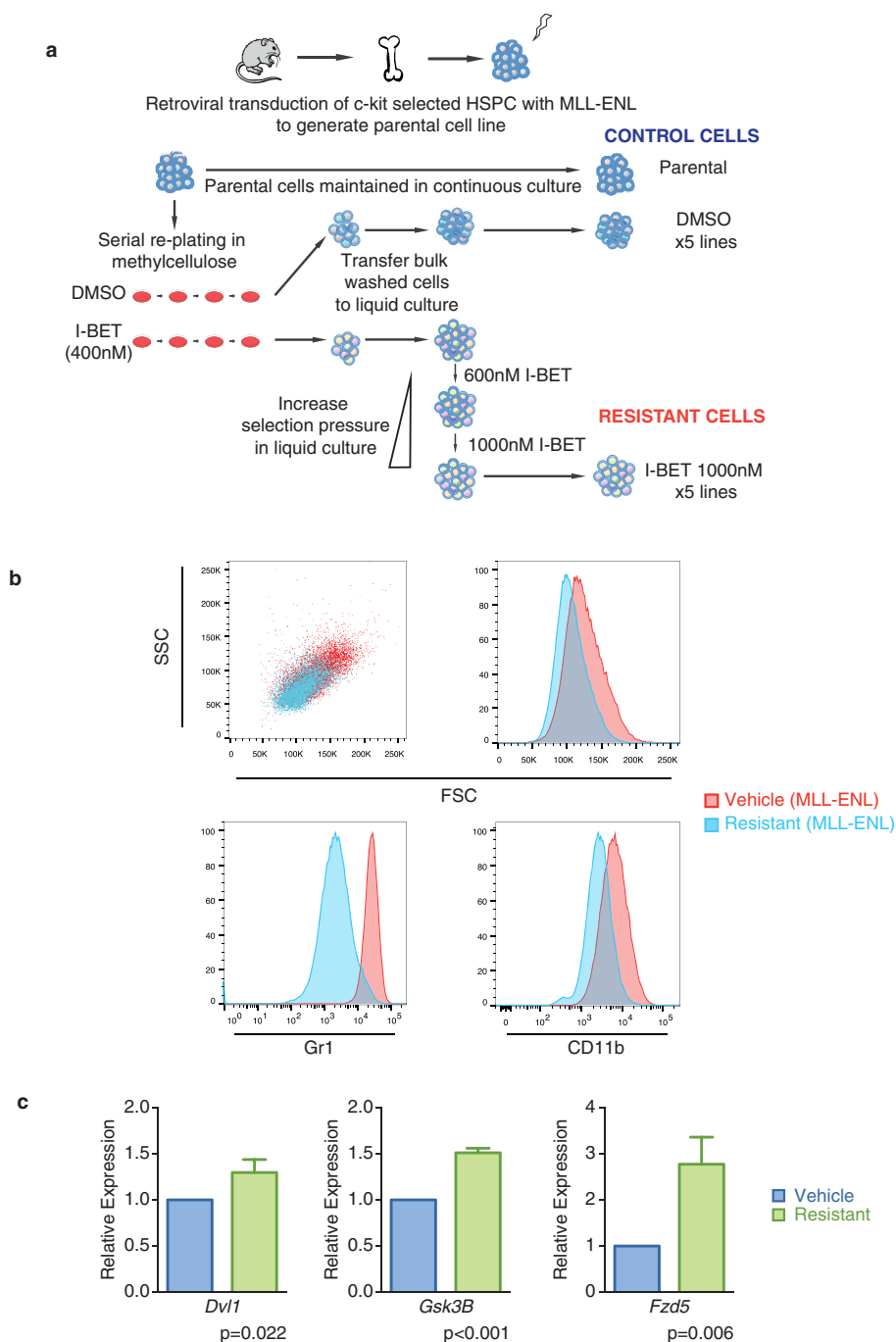
Patient material. Peripheral blood or bone marrow containing >80% blasts was obtained from patients following consent and under full ethical approval at each involved institute.

26. Zhao, W. *et al.* A new bliss independence model to analyze drug combination data. *J. Biomol. Screen.* **19**, 817–821 (2014).
27. Hu, Y. & Smyth, G. K. ELDA: extreme limiting dilution analysis for comparing depleted and enriched populations in stem cell and other assays. *J. Immunol. Methods* **347**, 70–78 (2009).
28. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).
29. Amarasinghe, K. C., Li, J. & Halgamuge, S. K. CoNVEX: copy number variation estimation in exome sequencing data using HMM. *BMC Bioinformatics* **14** (Suppl. 2), S2 (2013).
30. Koboldt, D. C. *et al.* VarScan 2: somatic mutation and copy number alteration discovery in cancer by exome sequencing. *Genome Res.* **22**, 568–576 (2012).
31. Cibulskis, K. *et al.* Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. *Nature Biotechnol.* **31**, 213–219 (2013).
32. McKenna, A. *et al.* The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* **20**, 1297–1303 (2010).
33. McLaren, W. *et al.* Deriving the consequences of genomic variants with the Ensembl API and SNP Effect Predictor. *Bioinformatics* **26**, 2069–2070 (2010).
34. Karolchik, D. *et al.* The UCSC Table Browser data retrieval tool. *Nucleic Acids Res.* **32**, D493–D496 (2004).
35. Zhang, Y. *et al.* Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* **9**, R137 (2008).
36. Tsigos, A., Haiminen, N., Bilal, E. & Utró, F. GenomicTools: a computational platform for developing high-throughput analytics in genomics. *Bioinformatics* **28**, 282–283 (2012).
37. Smyth, G. In *Bioinformatics and Computational Biology Solutions using R and Bioconductor* (eds Gentleman, R. *et al.*) 397–420 (Springer, 2005).
38. Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. B* **57**, 289–300 (1995).
39. Liao, Y., Smyth, G. K. & Shi, W. The Subread aligner: fast, accurate and scalable read mapping by seed-and-vote. *Nucleic Acids Res.* **41**, e108 (2013).
40. Liao, Y., Smyth, G. K. & Shi, W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**, 923–930 (2014).
41. Wu, D. *et al.* ROAST: rotation gene set tests for complex microarray experiments. *Bioinformatics* **26**, 2176–2182 (2010).
42. Subramanian, A. *et al.* Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl Acad. Sci. USA* **102**, 15545–15550 (2005).



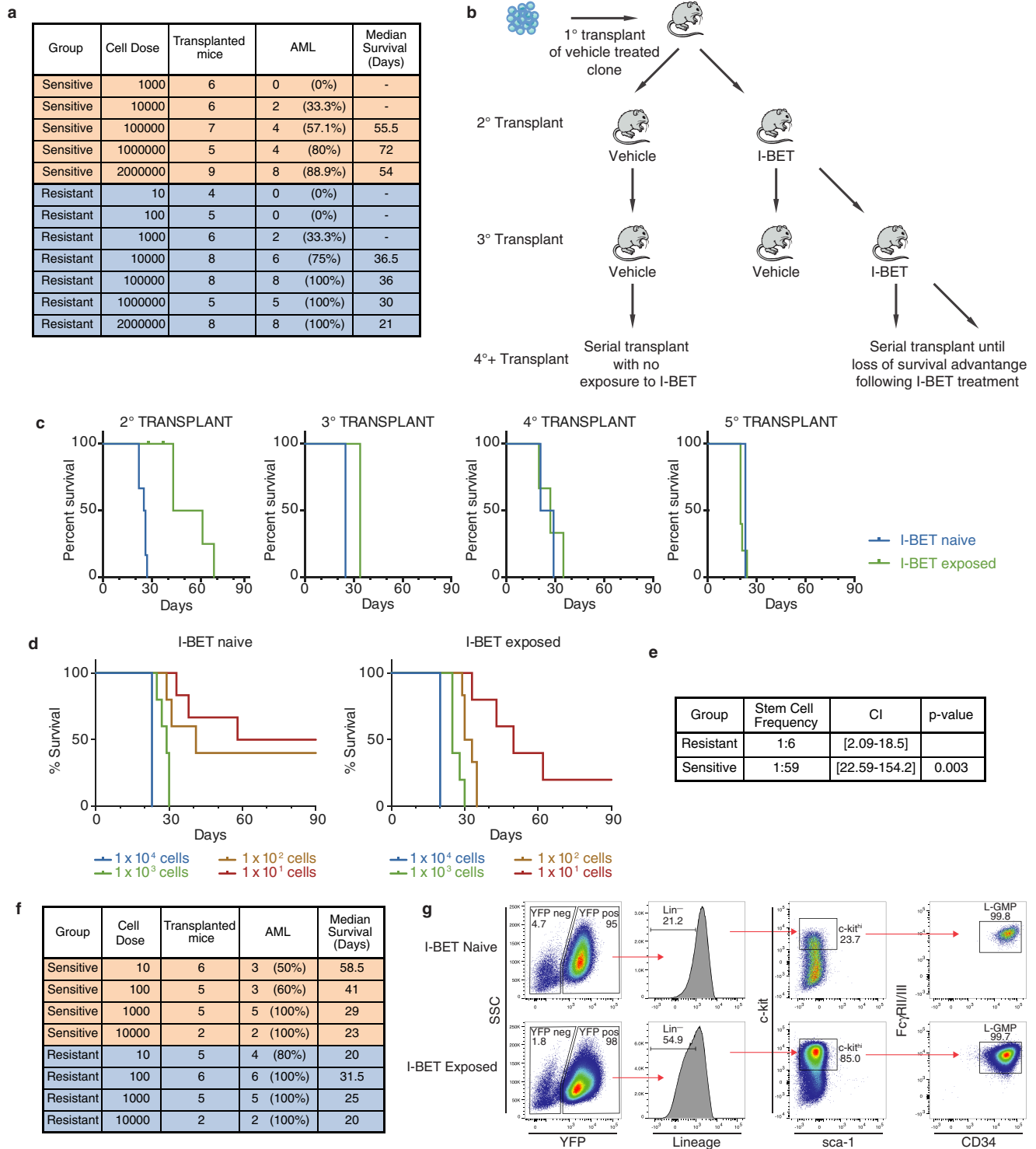
Extended Data Figure 1 | Establishment of a model of BET inhibitor resistance. **a**, Resistance to I-BET demonstrated in cell proliferation assays. Representative dose-response curve of a vehicle treated clone and a resistant clone stably maintained in 1,000 nM I-BET after 72 h of growth (mean \pm s.d., $n = 4$ per group). **b**, Cross-resistance to chemically distinct BET inhibitors demonstrated in cell proliferation assays. Representative dose-response curve of a vehicle-treated clone and a resistant clone after 72 h of growth (mean \pm s.d., $n = 4$ per group). **c**, Abrogation of response to I-BET-mediated cell cycle arrest. This is more evident in resistant clones stably maintained in higher concentrations of I-BET. Data from biological duplicate experiments (mean \pm s.e.m.). **d**, Representative flow plots of cell lines stably transduced with an inducible shRNA vector. Vector-positive cells constitutively express Venus. After the introduction of doxycycline, shRNA-expressing cells co-express Venus and dsRED. Selective disadvantage consequent to shRNA expression results in drop out of dsRED-positive cells from culture over time, which is assessed by flow cytometry. **e**, Independent inducible shRNAs specifically reduce the expression of *Brd4*, but not *Brd2* or *Brd3*, after 48–72 h of doxycycline. Messenger RNA levels in shRNA-positive cells normalized to

mRNA expression in shRNA-negative cells in biological duplicate experiments (mean \pm s.e.m.). **f**, *Brd4* protein levels are reduced in shRNA-positive cells. Uncropped blots are shown in Supplementary Fig. 1. **g**, In addition to resistance to *Brd4*, BET-inhibitor-resistant cells are also refractory to RNAi-mediated dual knockdown of *Brd3* and *Brd4*. shRNA-mediated knockdown of *Brd2* has minimal effect on both vehicle-treated and resistant clones. dsRED-positive cells normalized to day 1 after doxycycline exposure in biological duplicate experiments (mean \pm s.e.m.). **h**, Reduction of *Brd3/4* mRNA expression and *Brd2* mRNA expression with two independent shRNAs after 48–72 h of doxycycline. mRNA levels in shRNA-positive cells normalized to mRNA expression in shRNA-negative cells in biological duplicate experiments (mean \pm s.e.m.). **i**, Examination of vehicle-treated and resistant clones demonstrates no major morphological differences. **j**, Resistant clones are smaller and demonstrate homogeneity in size and complexity (FSC^{mid}/SSC^{low}) by flow cytometry. **k**, Resistant clones are enriched for L-GMPs (Lin⁻, Sca⁻, cKit⁺, CD34⁺, FcγR^{II/III}⁺). Representative FACS analysis of vehicle-treated and resistant clones, percentages represent proportion of parent gate.



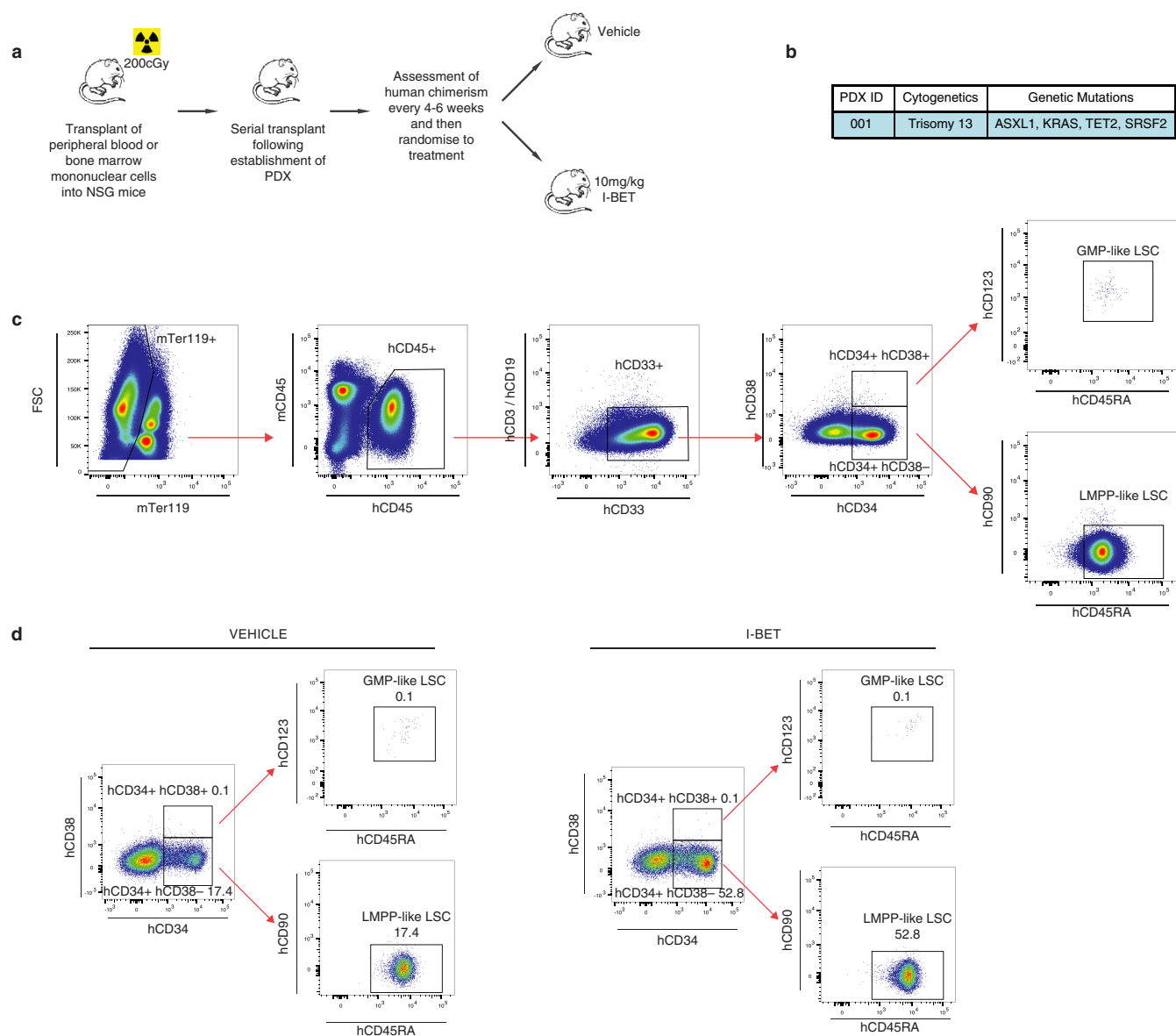
Extended Data Figure 2 | Resistance to BET inhibitors also arises from an immature cell compartment in MLL-ENL leukaemia. **a**, Strategy for the generation of resistant cell lines from primary HSPC after retroviral transduction with the oncogene MLL-ENL. The parental cell line was serially re-plated in cytokine supplemented semi-solid media containing either vehicle (0.1% DMSO) or 400 nM I-BET (\sim IC₄₀ of parental cell line). Cells in each plate were then washed and transferred to liquid culture to generate cell lines. Resistant cell lines were subsequently exposed to increasing selection pressure

in liquid culture. Vehicle-treated cell lines and the parental cell line were identically passed. **b**, Resistant cell lines bearing MLL-ENL are smaller and demonstrate homogeneity in size and complexity (FSC^{mid}/SSC^{low}) in addition to exhibiting an immature immunophenotype (Gr1⁺/CD11b⁺). Representative FACS analysis of vehicle-treated and resistant cell lines. **c**, Resistant cell lines bearing MLL-ENL demonstrate increased expression of WNT/ β -catenin pathway genes. qRT-PCR data performed in biological triplicate (mean \pm s.d.).



Extended Data Figure 3 | Resistance to BET inhibitors emerges from leukaemia stem cells. **a**, Transplant cohorts and survival of mice injected with vehicle-treated and resistant clones in limit dilution analyses of primary syngeneic transplants displayed in Fig. 2d, e. **b**, Experimental strategy for derivation of *in vivo* resistance to BET inhibitors in a MLL-AF9 leukaemia model. After primary transplant of a vehicle-treated clone, serial transplant of either I-BET-exposed or I-BET-naive leukaemias, derived from whole bone marrow of diseased mice, was undertaken until loss of I-BET-mediated survival advantage was observed. Treatment was started on days 11–13. **c**, Progressive loss of I-BET-mediated survival advantage observed in serial transplant generations. Kaplan–Meier curves of serial transplant generations. Second

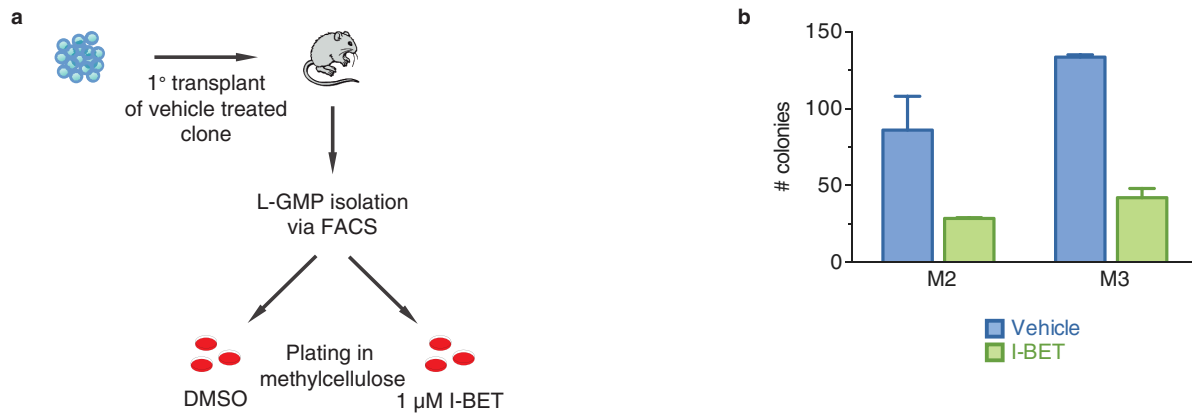
transplant: I-BET naive $n = 6$, I-BET exposed $n = 6$. Third transplant: I-BET naive $n = 2$, I-BET exposed $n = 3$. Fourth transplant: I-BET naive $n = 2$, I-BET exposed $n = 3$. Fifth transplant: I-BET naive $n = 4$, I-BET exposed $n = 5$. **d**, Limit dilution analyses of leukaemias derived from bone marrow of diseased mice chronically exposed to I-BET after the fourth transplant demonstrates that less than 10 cells are reliably able to transfer leukaemia. Kaplan–Meier curves of C57BL/6 mice injected with indicated number of cells. **e**, Chronic I-BET exposure significantly enriches for leukaemia stem cells *in vivo*. **f**, Transplant cohorts and survival of limit dilution analyses of data displayed in **d** and **e**. **g**, Gating strategy for identification of L-GMPs in whole mouse bone marrow.



Extended Data Figure 4 | Enrichment of a LMPP population in AML PDX.

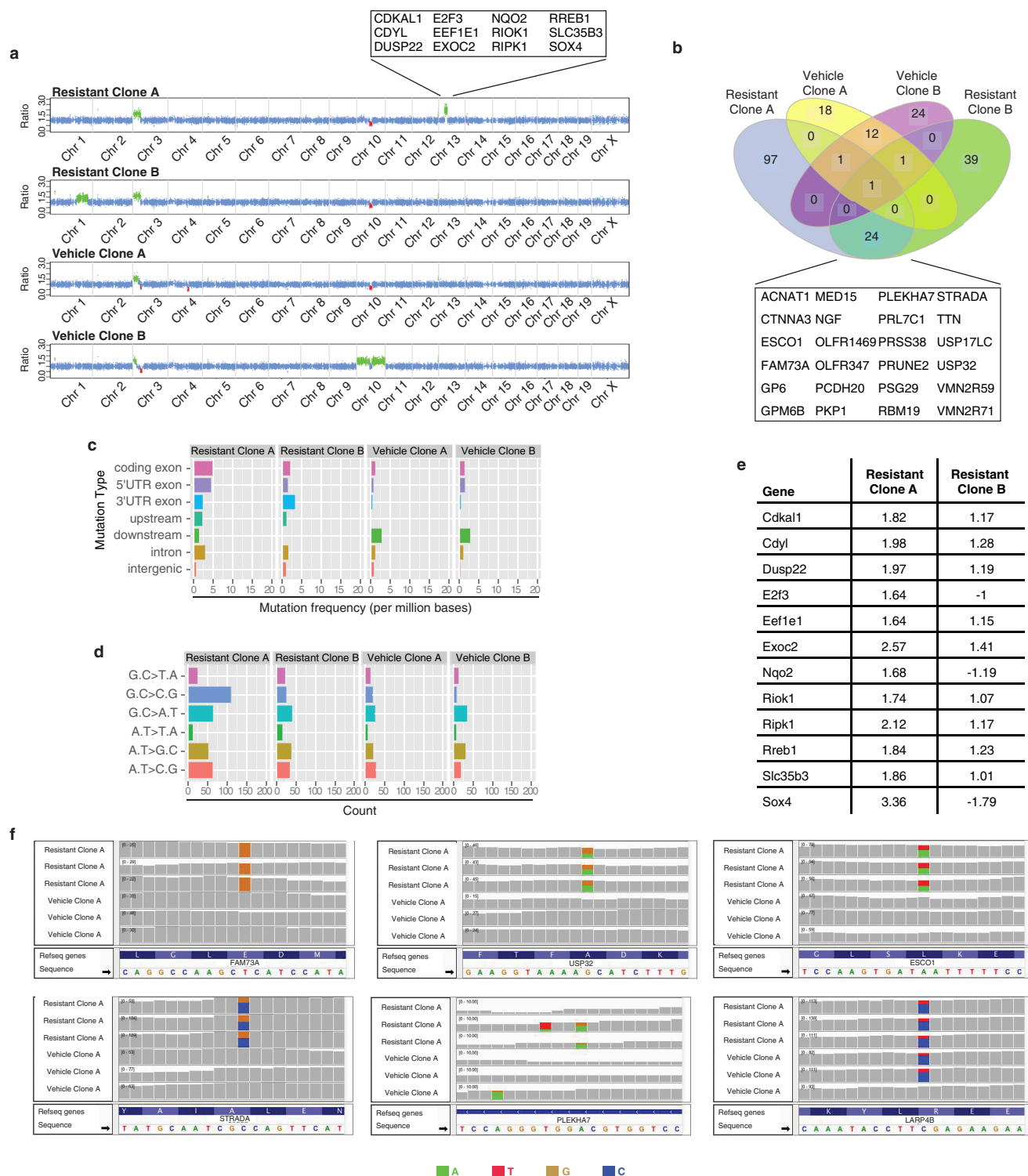
a, Experimental treatment strategy for treatment of NOD/SCID/*Il2rg*^{-/-} (NSG) mice bearing AML PDXs. Treated mice (with either vehicle or I-BET) belonged to identical transplant generations. **b**, Cytogenetic and genetic information of PDX models used. **c**, Gating strategy for identification of LMPP-like LSCs and GMP-like LSCs from mouse bone marrow. mTer119/mCD45 denotes mouse Ter119/CD45; hCD45/hCD33/hCD19/hCD33/hCD34/

hCD38/hCD123/hCD45RA/hCD90 denotes human CD45/CD3/CD19/CD33/CD34/CD38/CD123/CD45RA/CD90. **d**, Representative FACS analysis of bone marrow obtained from vehicle- and I-BET-treated mice demonstrating enrichment of LMPP-like LSCs in I-BET-treated mice. Events displayed are gated on mTer119⁻/hCD45⁺/hCD33⁺ cells and are expressed as a percentage of total hCD45⁺ cells.



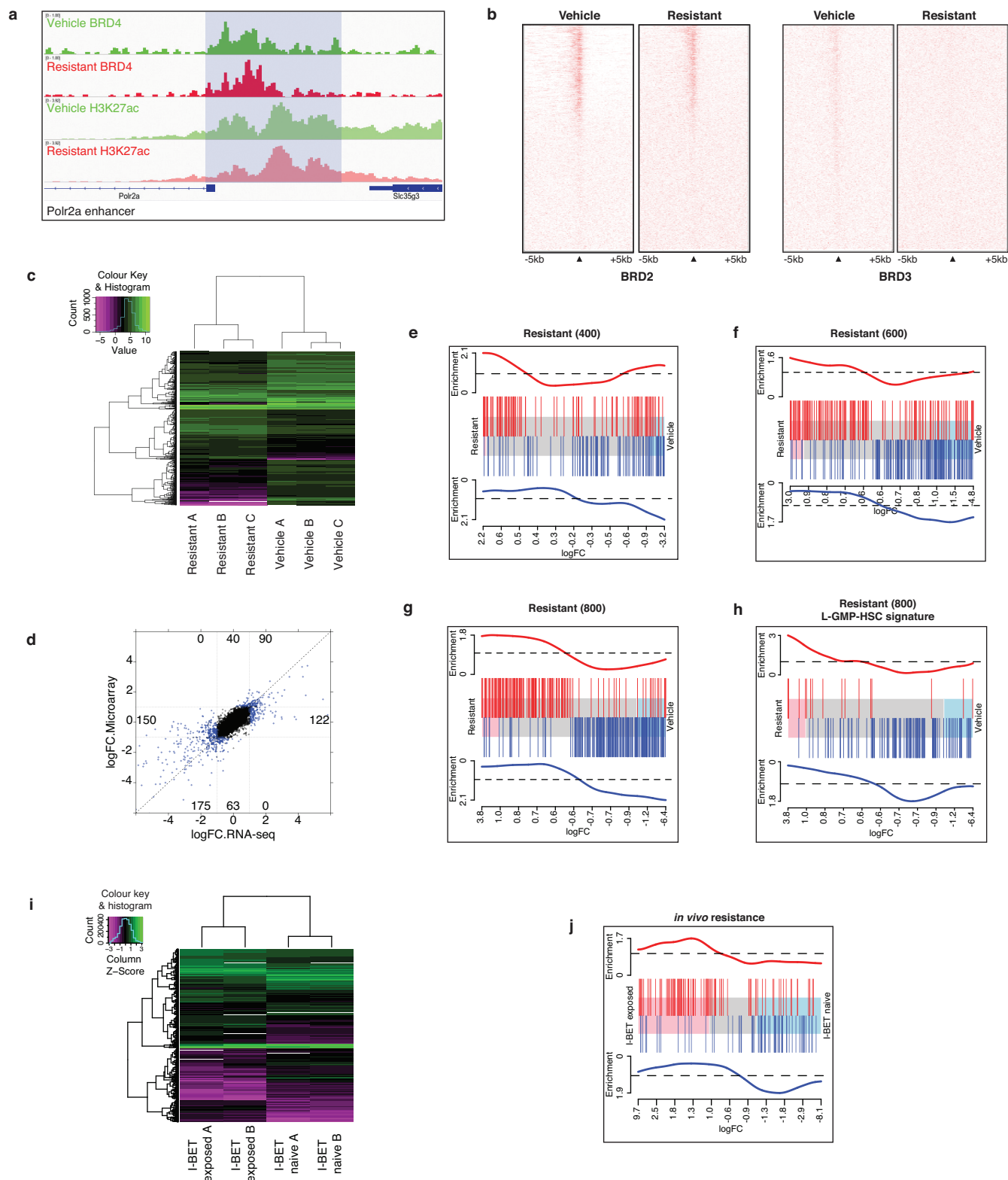
Extended Data Figure 5 | Intrinsic resistance to BET inhibition is not a feature of L-GMPs. **a**, Experimental strategy for testing intrinsic resistance of L-GMPs to BET inhibition. After syngeneic transplant of a vehicle-treated clone, L-GMPs were FACS-isolated from whole mouse bone marrow of diseased mice and cultured in cytokine supplemented semi-solid media

containing either vehicle (0.1% DMSO) or 1 μ M I-BET. **b**, L-GMPs do not demonstrate intrinsic resistance to I-BET. Colony counts after 7 days of growth in biological triplicate (see also Fig. 2h) experiments (mean \pm s.e.m.) of FACS-isolated L-GMPs after primary transplant of vehicle-treated clones in two additional independent mice. M2, mouse 2; M3, mouse 3.



Extended Data Figure 6 | Further genetic characterization of BET-inhibitor-resistant cells. **a**, Comparison of whole-exome sequencing (WES) data from early and late time points identifies non-advantageous passenger mutations. Data from WES of samples obtained at an earlier time point to that presented in Fig. 3d is shown. Call out box identifies genes within a small region on chromosome 13 in one resistant clone which demonstrate copy number gain and are associated with increased mRNA expression relative to non-resistant cells. **b**, Venn diagram demonstrating gene mutations shared between vehicle-treated and resistant clones. Highlighted in the call out box are 24 gene mutations shared between resistant clones but not found in

vehicle-treated clones. **c**, Resistant clones do not exhibit marked genetic instability with low mutation frequency observed. **d**, No specific mutation signature is identified in resistant clones. **e**, Correlation of genes identified in copy number gain region on chromosome 13 with gene expression data from the two resistant clones examined by WES. Fold change in gene expression compared to vehicle-treated clones obtained from microarray analysis is shown. **f**, Mutations detected by WES can be validated with data obtained from RNA sequencing (RNA-seq) of the same clones. Selected examples of mutations unique to resistant clones and shared between vehicle-treated and resistant clones is shown in integrative genomics viewer (IGV) tracks.



Extended Data Figure 7 | Further epigenetic and gene expression

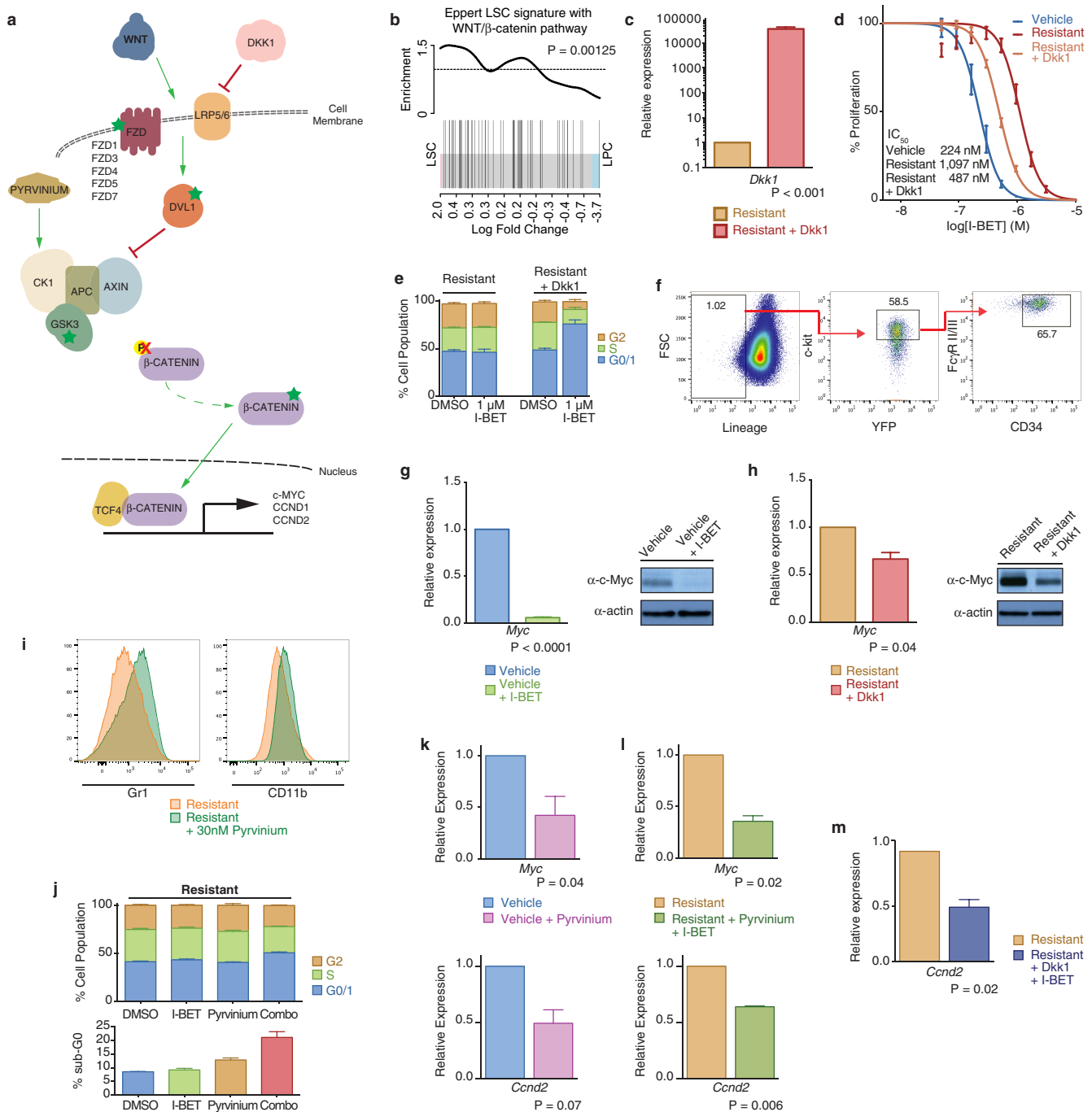
characterization of BET-inhibitor-resistant cells. **a**, BRD4 binding profile at *Polr2a* enhancer elements demonstrates no significant loss of BRD4 binding or H3K27ac levels in resistant clones. **b**, Genome wide profiling of BRD2 and BRD3 binding at TSSs comparing vehicle-treated and resistant clones is demonstrated in heat maps centred on the TSS of annotated genes with 5 kb flanking sequence either side. Red indicates higher density of reads in ChIP-seq data. **c**, Heat map of differential mRNA expression data from a vehicle-treated and resistant clone performed by RNA-seq in biological triplicate experiments. **d**, RNA-seq and microarray data are highly correlated.

Correlation of log₂ fold change (logFC) between RNA-seq and microarray data across all genes. No genes show opposing expression changes. Dotted line indicates $y = x$, blue dots represent genes that are significantly differentially expressed (gene expression log(FC) at least ± 1.0 , FDR corrected $P < 0.05$).

e–g, GSEA shows enrichment of LSC signature in I-BET-resistant cells, with resistant clones stably maintained in progressively higher concentrations of I-BET demonstrating increased enrichment of differentially expressed genes associated with a L-GMP self-renewal program. Barcode plot compares differential expression of genes in vehicle-treated and resistant clones to published microarray data comparing L-GMPs and MLL–AF9 cells propagated in liquid culture. Shaded area in the centre of plot shows genes ranked by fold change in expression in resistant relative to vehicle clones. Pink and blue shading represent significantly up- and downregulated genes, respectively.

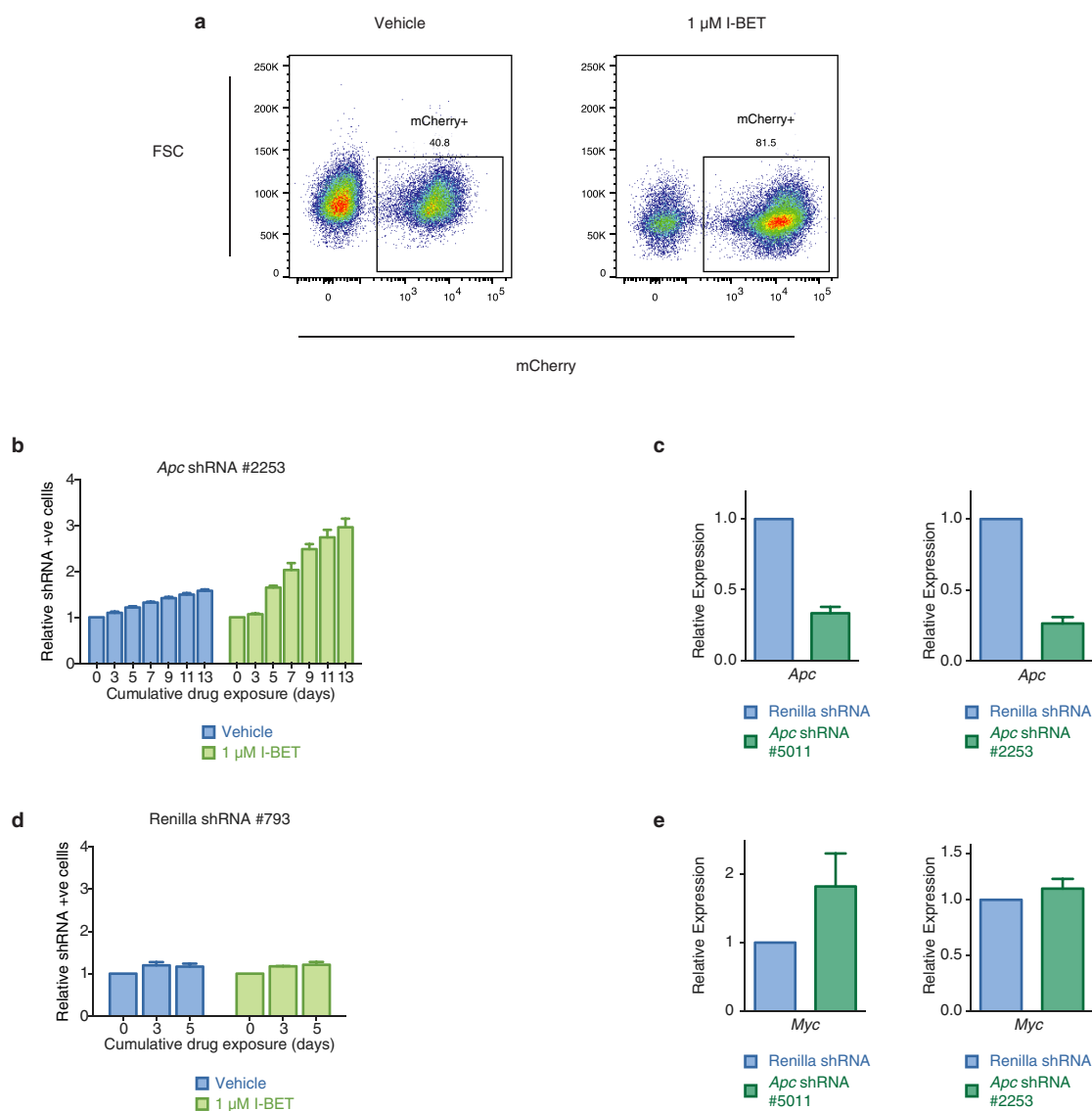
Upregulated and downregulated genes in the previously published LSC gene

expression signature are shown in red and blue, respectively. Resistant (400) upregulated FDR = 1.2×10^{-1} , downregulated FDR = 9.3×10^{-3} . Resistant (600) upregulated FDR $< 1.0 \times 10^{-4}$, downregulated FDR $< 2.5 \times 10^{-4}$. Resistant (800) upregulated FDR $< 5.0 \times 10^{-5}$, downregulated FDR $< 5.0 \times 10^{-5}$. **h**, GSEA demonstrates that resistant clones show significant enrichment for genes associated with a self-renewal program identified from L-GMPs arising from haematopoietic stem cells (L-GMP HSCs). Upregulated FDR = 4.69×10^{-2} , downregulated FDR = 1.3×10^{-3} . **i**, RNA-seq identifies enrichment of LSC gene expression signature following chronic *in vivo* BET inhibitor exposure. Heat map of differential mRNA expression data from RNA-seq of leukaemias from the bone marrow of I-BET-exposed ($n = 2$) and I-BET-naive ($n = 2$) mice after the fourth transplant. **j**, GSEA of RNA-seq data identifies enrichment of a previously published LSC gene expression signature in leukaemias chronically exposed to I-BET *in vivo*. Barcode plot compares differential expression of genes in I-BET-exposed and I-BET-naive leukaemias to published data comparing L-GMPs and MLL–AF9 cells propagated in liquid culture. Shaded area in the centre of plot shows genes ranked by fold change in expression in I-BET-exposed relative to I-BET-naive leukaemias. Pink and blue shading represent significantly up- and downregulated genes, respectively. Upregulated and downregulated genes in the previously published LSC gene expression signature are shown in red and blue, respectively, and correlate with expression of genes in the I-BET-exposed leukaemias (FDR = 0.05).



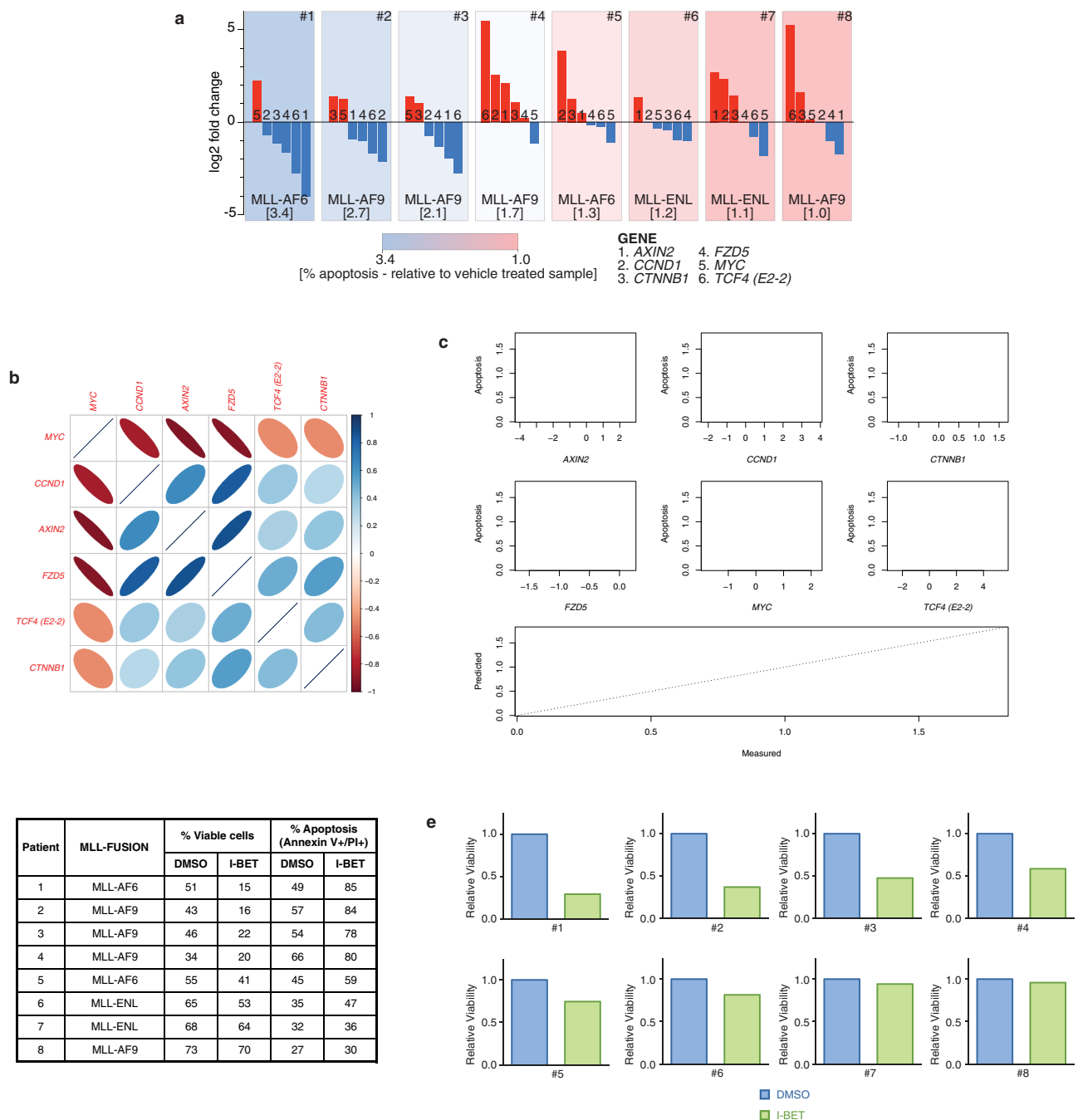
Extended Data Figure 8 | Negative regulation of Wnt/β-catenin signalling in resistant clones re-establishes sensitivity to BET inhibition. **a**, Schematic representation of the Wnt/β-catenin pathway. Highlighted by green stars are components of the pathway identified from transcriptome data which are significantly upregulated (>1.5-fold change, FDR < 0.05) in resistant clones relative to vehicle-treated clones. **b**, GSEA of previously published human LSC gene expression data demonstrates enrichment of the WNT/β-catenin pathway. **c**, *Dkk1* expression in the resistant cells before and after retroviral transduction of mouse *Dkk1*. qRT-PCR data from biological triplicate experiments (mean ± s.d.). **d**, Partial restoration of sensitivity to BET inhibition is observed in resistant clones after transduction with *Dkk1*. Dose-response curve of a vehicle-treated clone and resistant clone with and without expression of *Dkk1* after 72 h of growth (mean ± s.e.m., *n* = 16 per group). **e**, Restoration of BET inhibitor induced cell-cycle arrest in resistant clones stably transduced with *Dkk1*. Flow cytometric analysis after 48 h exposure to either vehicle or 1,000 nM I-BET in biological triplicate experiments (mean ± s.e.m.). **f**, Resistant clones stably expressing *Dkk1* do not show immunophenotypic enrichment for L-GMPs (see Extended Data Fig. 1k for comparison). Representative FACS analysis of resistant clone expressing *Dkk1*,

percentages represent proportion of parent gate. **g**, Abrogation of *Myc* mRNA and protein expression in vehicle-treated clones after treatment with I-BET. qRT-PCR data of *Myc* expression in a vehicle-treated clone after 6 h of treatment with 1 μM I-BET151 in biological triplicate experiments (mean ± s.d.). Uncropped blots are found in Supplementary Fig. 1. **h**, Negative regulation of Wnt/β-catenin signalling by *Dkk1* in resistant clones results in decreased expression of *Myc*. qRT-PCR data from biological triplicate experiments (mean ± s.d.). **i**, Small molecule inhibition of Wnt/β-catenin pathway expression re-establishes sensitivity to BET inhibition. Exposure of resistant clones to the Wnt/β-catenin pathway inhibitor pyriminium also results in re-expression of Gr1⁺ and CD11b⁺. Representative FACS analysis of resistant clone in the presence or absence of pyriminium. **j**, Pyriminium synergises with I-BET to induce a modest cell cycle arrest and an induction of cell death (sub-G0 cell fraction). Data from biological triplicate experiments (mean ± s.e.m.). **k**, **l**, Pyriminium reduces the expression of Wnt/β-catenin target genes such as *Myc* and *Ccnd2* in vehicle-treated and resistant cells. qRT-PCR data from biological duplicate experiments (mean ± s.e.m.). **m**, These findings are similar to those seen for resistant cells stably expressing *Dkk1*.



Extended Data Figure 9 | shRNA-mediated knockdown of *Apc* confers resistance to sensitive clones **a**, shRNA-mediated knockdown of *Apc*, a negative regulator of Wnt/ β -catenin signalling, confers resistance to vehicle-treated clones. BET inhibitor treatment enriches for shRNA-containing (mCherry-positive) cells. Representative FACS plots after 7 days of cumulative drug exposure to either vehicle (0.1% DMSO) or 1 μ M I-BET in a vehicle clone transduced with an *Apc* shRNA. **b**, **c**, Independent shRNAs directed against *Apc* confer resistance to vehicle-treated clones. Viable, shRNA-positive cells after treatment with either vehicle or I-BET normalized to day 0 performed

in biological triplicate (mean \pm s.d.). qRT-PCR data from FACS-isolated shRNA-containing cells performed in biological duplicate (mean \pm s.e.m.). **d**, I-BET treatment of vehicle-treated clones transduced with a non-targeting shRNA does not enrich for shRNA-containing cells. Viable, shRNA-positive cells after treatment with either vehicle or I-BET normalized to day 0 performed in biological triplicate (mean \pm s.e.m.). **e**, shRNA-mediated knockdown of *Apc* results in increased expression of Wnt/ β -catenin target gene *Myc*. qRT-PCR data from FACS isolated shRNA containing cells performed in biological duplicate (mean \pm s.e.m.).



Extended Data Figure 10 | WNT/β-catenin pathway expression correlates with responsiveness to I-BET in primary human AML samples. **a**, Assessment of β-catenin pathway gene expression in eight primary human AML samples with associated response to I-BET exposure. Each panel represents an individual primary human AML sample, with genetic abnormality denoted. Waterfall plot of relative qRT-PCR expression data of key β-catenin pathway genes (*AXIN2*, *CCND1*, *CTNNB1*, *FZD5*, *MYC*, *TCF4* (also known as *E2-2*)) is displayed. Each bar is labelled 1–6 according to gene represented. Relative apoptosis observed after 48 h exposure to 500 nM I-BET versus vehicle (0.1% DMSO) is denoted in square parenthesis and is also represented as a heat map background shading in each panel. **b**, log₂-transformed expression levels of selected genes in the WNT/β-catenin pathway were measured using

qRT-PCR. A corrgram shows the genes are highly correlated with each other. The colour and thinness of the ellipse indicate the strength of correlation (a line is perfect correlation; a circle is uncorrelated). The ellipse direction indicates the sign of the correlation (correlated: right/blue, inversely correlated: left/red). **c**, Expression of selected genes is correlated with apoptosis. Scatterplots show apoptosis versus the log₂ expression level of each gene. Expression of five genes (*CCND1*, *CTNNB1*, *FZD5*, *MYC* and *TCF4*) predicts apoptosis. The relationship is highlighted in a plot of apoptosis predicted using a multiple linear regression model with the five genes versus the actual data. **d**, Apoptosis observed after 48 h exposure to either vehicle (0.1% DMSO) or 500 nM I-BET across eight primary human AML samples. **e**, Relative viability of primary human AML samples after treatment with I-BET.

Transcriptional plasticity promotes primary and acquired resistance to BET inhibition

Philipp Rathert^{1*}, Mareike Roth^{1*}, Tobias Neumann¹, Felix Muerdter¹, Jae-Seok Roe², Matthias Muhar¹, Sumit Deswal¹, Sabine Cerny-Reiterer^{3,4}, Barbara Peter^{3,4}, Julian Jude¹, Thomas Hoffmann¹, Lukasz M. Boryn¹, Elin Axelsson¹, Norbert Schweifer⁵, Ulrike Tontsch-Grunt⁵, Lukas E. Dow⁶, Davide Gianni⁵, Mark Pearson⁵, Peter Valent^{3,4}, Alexander Stark¹, Norbert Kraut⁵, Christopher R. Vakoc² & Johannes Zuber¹

Following the discovery of BRD4 as a non-oncogene addiction target in acute myeloid leukaemia (AML)^{1,2}, bromodomain and extra-terminal protein (BET) inhibitors are being explored as a promising therapeutic avenue in numerous cancers^{3–5}. While clinical trials have reported single-agent activity in advanced haematological malignancies⁶, mechanisms determining the response to BET inhibition remain poorly understood. To identify factors involved in primary and acquired BET resistance in leukaemia, here we perform a chromatin-focused RNAi screen in a sensitive MLL–AF9;Nras^{G12D}-driven AML mouse model, and investigate dynamic transcriptional profiles in sensitive and resistant mouse and human leukaemias. Our screen shows that suppression of the PRC2 complex, contrary to effects in other contexts, promotes BET inhibitor resistance in AML. PRC2 suppression does not directly affect the regulation of Brd4-dependent transcripts, but facilitates the remodelling of regulatory pathways that restore the transcription of key targets such as *Myc*. Similarly, while BET inhibition triggers acute *MYC* repression in human leukaemias regardless of their sensitivity, resistant leukaemias are uniformly characterized by their ability to rapidly restore *MYC* transcription. This process involves the activation and recruitment of WNT signalling components, which compensate for the loss of BRD4 and drive resistance in various cancer models. Dynamic chromatin immunoprecipitation sequencing and self-transcribing active regulatory region sequencing of enhancer profiles reveal that BET-resistant states are characterized by remodelled regulatory landscapes, involving the activation of a focal *MYC* enhancer that recruits WNT machinery in response to BET inhibition. Together, our results identify and validate WNT signalling as a driver and candidate biomarker of primary and acquired BET resistance in leukaemia, and implicate the rewiring of transcriptional programs as an important mechanism promoting resistance to BET inhibitors and, potentially, other chromatin-targeted therapies.

BRD4 is a chromatin reader that regulates transcription through linking histone acetylation and core components of the transcriptional apparatus⁷. Recent studies suggest that BRD4 regulates distinct gene sets through interacting with context-specific enhancers and transcription factors^{8,9}. However, the mechanisms underlying the wide range of sensitivity to BET inhibition remain elusive, and so far no predictive biomarker has been identified. Towards understanding these mechanisms, we sought to functionally identify chromatin factors that are required for rendering AML cells sensitive to JQ1, a well-known BET inhibitor¹⁰. To this end, we constructed a microRNA-embedded short hairpin RNA (shRNA_{mir}) library covering 626 chromatin regulators and screened it in the same MLL–AF9;Nras^{G12D}-driven model that led to the identification of BRD4 as a candidate target¹ (Fig. 1a). Deep-sequencing following transduction (T0) and 7 days of selection (T1)

identified chromatin-associated dependencies, including Smarca4 and Brd4 as top hits (Extended Data Fig. 1a–c). To control for unspecific events, we mixed the GFP⁺ library population with mCherry⁺ control cells, and subsequently treated with DMSO or JQ1. While mCherry⁺ cells disappeared over time (Fig. 1b), GFP⁺ cells survived and eventually grew in the presence of 50 nM JQ1 (corresponding to an IC₇₀ dose; Extended Data Fig. 1d), indicating that resistance emerged from shRNA-mediated effects. Four shRNAs showed an outstanding enrichment that was consistent between replicates, despite almost 5 weeks of independent culture (Fig. 1c and Extended Data Fig. 1a, e). These included two shRNAs targeting Suz12, one targeting Psp1, and one previously characterized potent Dnmt3a shRNA¹¹. All four shRNAs strongly suppressed their target mRNA (Extended Data Fig. 1f), and validated to promote JQ1 resistance in single assays (Fig. 1d and Extended Data Fig. 2a).

The finding that suppression of Suz12, a component of the PRC2 complex, promotes resistance to JQ1 was surprising in two ways. First, a recent report, based on studies in nerve sheath tumours, has implicated Suz12 deficiency as a condition that increases sensitivity to BET inhibition¹². Second, several studies (including work from our group) have characterized PRC2 as a requirement in MLL–AF9-driven AML^{13,14}. Notably, the most potent Suz12 shRNA in previous studies (Suz12.1676) did not score in the pooled screen and validated to strongly inhibit proliferation in our model (Fig. 1e). However, when we added JQ1, Suz12.1676-expressing cells were rapidly enriched, indicating that Suz12 deficiency turns from a detrimental into a favourable condition. Similar effects were observed using potent shRNAs targeting Ezh2 and Eed, two other PRC2 components (Fig. 1e and Extended Data Fig. 2b). We also validated this phenomenon using Tet-regulatable RNA interference, where we included a validated *Myc* shRNA¹⁵ to rule out that resistance is merely a consequence of reduced proliferation (Extended Data Fig. 2c). Resistant cells generated through Suz12 suppression showed a global loss of H3K27me3 (Fig. 1f), and were also refractory to the effects of JQ1 in methyl-cellulose assays (Extended Data Fig. 2d) and *in vivo* (Fig. 1g). While recipient mice of Suz12-deficient cells, consistent with previous observations¹⁴, showed a delayed disease progression, JQ1 had no anti-leukaemic effects in this context (Fig. 1g). Together, these data demonstrate that loss of Suz12 impairs rather than promotes BET sensitivity in AML, revealing another scenario where PRC2 plays opposing roles in different cancers.

To search for underlying mechanisms, we investigated whether our hits affected the regulation of BRD4-dependent transcripts, and found that the acute response to JQ1 is largely unchanged in the absence of Suz12, Dnmt3a or Psp1 (Fig. 2a). However, under long-term treatment, 60–80% of JQ1-induced changes reverted and transcription of the key target *Myc*^{1,3} was restored (Fig. 2a–c and Extended Data Fig. 2e).

¹Research Institute of Molecular Pathology (IMP), Vienna Biocenter (VBC), 1030 Vienna, Austria. ²Cold Spring Harbor Laboratory, Cold Spring Harbor, New York 11724, USA. ³Department of Internal Medicine I, Division of Hematology and Hemostaseology, Medical University of Vienna, 1090 Vienna, Austria. ⁴Ludwig Boltzmann Cluster Oncology, Medical University of Vienna, 1090 Vienna, Austria. ⁵Boehringer Ingelheim – Regional Center Vienna GmbH, 1121 Vienna, Austria. ⁶Department of Medicine, Hematology & Medical Oncology, Weill Cornell Medical College, New York 10065, USA.

*These authors contributed equally to this work.

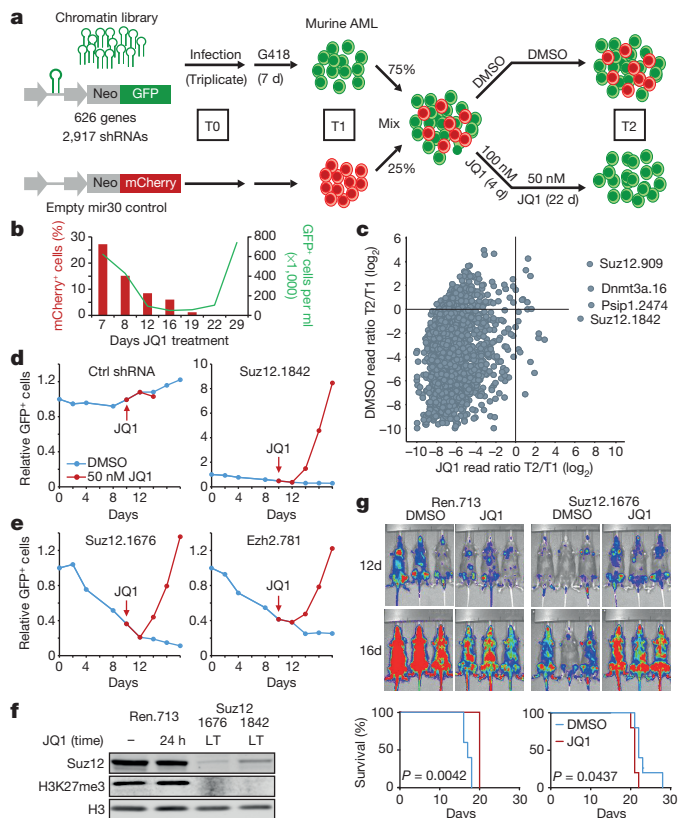


Figure 1 | Multiplexed shRNA screening identifies chromatin factors that prevent resistance to BET inhibition. **a**, Schematic of the multiplexed screening strategy. Mouse MLL-AF9/Nras^{G12D} AML cells were infected with a library targeting 626 chromatin-associated genes (GFP⁺) or empty control (mCherry⁺). G418-selected cell populations were mixed and treated with DMSO or 100 nM JQ1 for 4 days followed by 50 nM JQ1 for 22 days. Genomic DNA isolated from T0, T1 and T2 was used to amplify and deep-sequence shRNA guides. **b**, Relative abundance of mCherry⁺ control cells and absolute number of GFP⁺/shRNA-expressing cells in the screen population over time. **c**, Scatter plot showing the average ratio of normalized reads before (T1) and after 26 days of DMSO- or JQ1 treatment (T2) for all 2,917 shRNAs. **d, e**, Competitive proliferation assays of MLL-AF9/Nras^{G12D} leukaemia cells expressing the indicated shRNAs. Shown is the relative fraction of GFP⁺/shRNA⁺ cells relative to the initial measurement. After 10 days, each sample was split in half, treated with DMSO or 50 nM JQ1 and analysed over 8 days. **f**, Immunoblotting of Suz12 and H3K27me3 in resistant AML cells expressing the indicated Suz12 shRNAs (LT, long-term culture in 50 nM JQ1 for 6 weeks) and Ren.713 control cells (with or without treatment with 200 nM JQ1 for 24 h). **g**, Top, bioluminescent imaging of mice transplanted with 1×10^5 MLL-AF9/Nras^{G12D} leukaemia cells expressing the indicated shRNAs. Treatment with JQ1 (50 mg kg⁻¹ per day) or DMSO carrier was initiated at day 3 after transplantation. Bottom, Kaplan-Meier survival curves of control and JQ1-treated mice ($n = 5$). Statistical significance was calculated using a log-rank test.

To investigate this rebound phenomenon, we focused subsequent analyses on PRC2, a writer of repressive H3K27me3 marks¹⁶. Based on recent reports showing that BRD4 interacts with WHSC1L1¹⁷, a writer of H3K36 methylation marks that can recruit PRC2¹⁸, we wondered whether PRC2 is involved in repressing BRD4 targets following BET inhibition. However, chromatin immunoprecipitation sequencing (ChIP-seq) analysis before and after JQ1 treatment showed no increase in H3K27me3 at Brd4 occupancy sites (Extended Data Fig. 2f, g), which together with RNA-seq studies suggests that PRC2 does not regulate these genes directly.

Next, we investigated whether the rebound of BRD4 targets is associated with changes in the enhancer landscape. Under long-term JQ1 treatment, resistant cells showed a global loss of enhancer-associated H3K27ac marks, which was reversible upon drug withdrawal (Extended Data Fig. 3a). At the level of *Myc*, the loss of H3K27ac

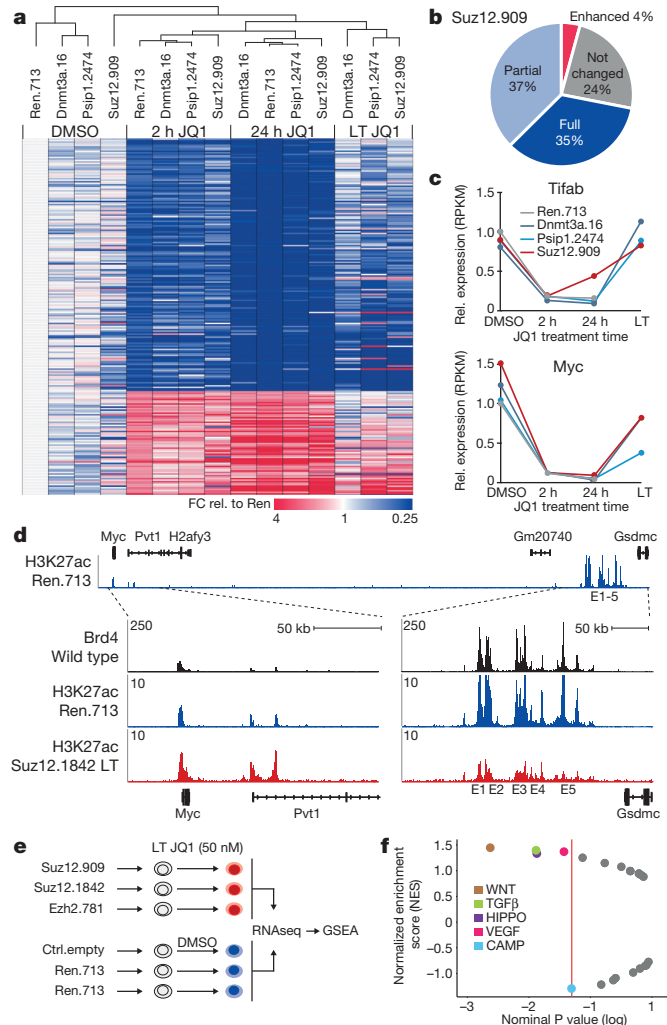


Figure 2 | BET-resistant AML cells restore the transcription of key Brd4 target genes through remodelling of regulatory landscapes and pathways. **a**, Heat map showing RPKM fold change (FC) of 235 response genes in MLL-AF9/Nras^{G12D} leukaemia cells expressing the indicated shRNAs after treatment with DMSO or JQ1 (200 nM) for 2 h or 24 h, and after long-term (LT) JQ1 treatment (50 nM) for 6 weeks. Response genes were defined based on JQ1-induced changes in Ren.713 expressing leukaemia cells (FC >2/<0.5 after 2 h and FC >2/<0.33 after 24 h). Samples are presented according to non-hierarchical clustering; genes are grouped in up- or downregulated targets in Ren.713 control cells, and sorted by their average re-expression in resistant leukaemia. **b**, Response genes in **a** were grouped into four categories based on the divergence of their expression in resistant AML compared to AML expressing Ren.713 (see legend of Extended Data Fig. 2e for details). **c**, Time course of *Tifab* and *Myc* transcript levels in MLL-AF9/Nras^{G12D} leukaemia expressing the indicated shRNAs, relative to untreated cells expressing Ren.713. **d**, ChIP-seq occupancy profiles of Brd4 and H3K27ac in *Myc* regulatory regions in sensitive AML cells (wild type and Ren.713) and Suz12.1842-expressing resistant AML cells under long-term treatment with JQ1 (50 nM). The H3K27ac profile in Ren.713 control AML cells is shown on top for the entire region (~2 Mb), together with validated transcript models from the mm10 genome assembly. Additional tracks are zoomed in on the proximal region and a distal region containing an established cluster of enhancers (E1-E5)¹⁹, as indicated (dotted lines). The y axis reflects the number of normalized cumulative tag counts in each region. **e**, Schematic of RNA-seq profiling in resistant AML cells generated through expression of three independent shRNAs targeting Suz12 or Ezh2, which are treated as biological triplicates and compared to triplicate control cells. **f**, Gene set enrichment analysis of expression changes in 22 KEGG signalling pathways in resistant AML cells generated as described in **e**. Plotted are normalized enrichment scores (NES) against nominal *P* values; significantly altered gene sets ($P < 0.05$) are indicated in the graph.

was most prominent at a 3' cluster of lineage-specific enhancers known to be required for *Myc* transcription¹⁹ (Fig. 2d). In parallel, we observed a strong focal gain of H3K27ac in the first intron of *Pvt1* (a long non-coding RNA 3' of *Myc*), which was the tenth most prominent of only 119 genomic regions that gained H3K27ac in resistant cells (Fig. 2d and Extended Data Fig. 3b, c). Together, these findings reveal that BET resistance is associated with profound changes in the enhancer landscape.

To probe regulatory pathways involved in the re-expression of Brd4 targets, we analysed the transcriptomes of sensitive and resistant cells generated using three independent PRC2 shRNAs (Fig. 2e). Gene set enrichment analysis²⁰ identified alterations in five major signalling pathways (Fig. 2f). Almost half of the genes associated with upregulated pathways contain H3K27me3 marks in sensitive cells (Extended Data Fig. 3d–f), indicating that loss of PRC2 can facilitate their transcriptional activation. The most significant alteration was an upregulation of genes associated with Wnt signalling, which can drive *Myc* transcription²¹ and has important functions in normal and leukaemic stem cells (LSC)^{22,23}. Of note, resistant cells in our model were not enriched for LSC-associated surface markers and expression signatures (Extended Data Fig. 4a–c), which does not exclude a role of LSC in resistance. Collectively, our findings suggest that loss of PRC2 promotes resistance through facilitating the derepression of compensatory pathways that restore the transcription of critical BRD4 target genes.

Given this complexity, we sought to explore whether any phenomena in our model of acquired resistance are also associated with primary BET resistance. To this end, we determined the JQ1 sensitivity in 246 human cell lines (Extended Data Fig. 5a), and selected three sensitive and three resistant lines from three cancer subtypes for dynamic transcriptional profiling. While RNA-seq profiles distinguished cancer subtypes as expected, sensitive and resistant contexts could not be differentiated, neither through their steady-state transcriptomes, nor through comparing JQ1-induced changes (Extended Data Fig. 5b, c). In fact, BET inhibition triggered highly distinct responses regardless of whether *MYC* is affected, and even cancers of similar tissue and sensitivity profile do not share greater overlaps (Fig. 3a and Extended Data Fig. 5d–f). To test whether this heterogeneity is associated with differential functions of JQ1 targets, we used Tet-regulated shRNAmirs to profile BRD2/3/4-dependent transcripts in sensitive and resistant leukaemia cells. In both cases, JQ1-induced changes showed the largest overlap with BRD4-knockdown profiles (Extended Data Fig. 6a–d), indicating that suppression of BRD4 dominates the effects of BET inhibition in leukaemia.

Despite this heterogeneity, a closer look at the few commonly regulated genes (Extended Data Fig. 6e) revealed two interesting findings: (1) one of two transcripts commonly induced by JQ1 was *HEXIM1*, which we found upregulated in all analysed cancers (Extended Data Fig. 6f), indicating that it could serve as a pharmacodynamic biomarker. (2) Surprisingly, one of three transcripts repressed after 2 h of JQ1 treatment in both sensitive and resistant leukaemias was *MYC*. In sensitive leukaemias, this repression was durable and led to strong suppression of the *MYC* protein (Fig. 3b), while all resistant leukaemias showed a rapid rebound of *MYC* transcription and, consequently, only minimal protein suppression. Similar effects were observed in a larger panel of haematopoietic cell lines (Fig. 3c and Extended Data Fig. 6g), indicating that primary BET resistance is not due to BRD4-independent regulation of critical targets, but driven by compensatory mechanisms that rapidly restore their transcription.

To investigate how this rebound phenomenon is encoded in the enhancer landscape, we performed H3K27ac ChIP-seq, which revealed that resistant K-562 cells contain several H3K27ac peaks around *MYC* that are missing in sensitive MOLM-13 cells. These include a strong occupancy in the first intron of *PVT1* (Fig. 3d), the same region that gained H3K27ac in our mouse model (Fig. 2d). To systematically analyse how these putative *MYC* enhancers change their activity upon BET

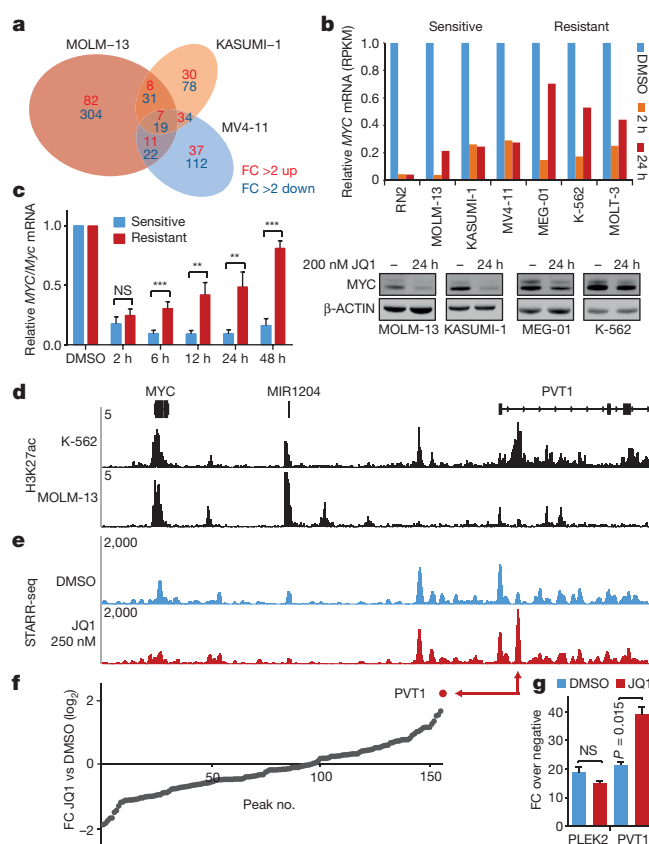


Figure 3 | Dynamic transcriptional and enhancer profiling of sensitive and resistant cancer cell lines. **a**, Venn diagram depicting the overlap of expression changes in three sensitive AML lines after 2 h of JQ1 (200 nM). **b**, Top, *MYC* mRNA levels in indicated cell lines after 2 h or 24 h of JQ1 treatment (200 nM), relative to DMSO-treated cells. Bottom, *MYC* immunoblotting in indicated cell lines before and after 24 h of JQ1. **c**, *MYC* mRNA levels across six sensitive and six resistant cell lines at indicated time points following JQ1 (200 nM), normalized to DMSO (mean \pm s.e.m.; Student's *t*-test; ***P* = 0.01; ****P* < 0.01). **d**, ChIP-seq occupancy profiles of H3K27ac at the *MYC* locus in K-562 and MOLM-13 cells. Validated transcript models from the hg19 genome assembly are shown above; y axes reflect normalized cumulative tag counts in each region. **e**, STARR-seq fragment densities in K-562 cells treated with DMSO or 250 nM JQ1. Read densities are shown as unique reads per million. **f**, Sorted ratios of STARR-seq signals in DMSO versus 250 nM JQ1-treated cells for all 156 identified STARR-seq peaks. **g**, Luciferase reporter assay measuring the enhancer activity of the *PVT1* element and an unrelated enhancer (*PLEK2*) on a minimal *MYC* promoter. Shown are fold changes of normalized luciferase signal over background \pm 250 nM JQ1 (*n* = 3; mean \pm s.e.m., Student's *t*-test with Welch's correction).

inhibition, we performed self-transcribing active regulatory region sequencing (STARR-seq; a new method for high-throughput functional enhancer analysis²⁴) using a library covering 3.1 megabase pairs surrounding *MYC* (Extended Data Fig. 7a). While the *PVT1* element had little activity in DMSO-treated K-562 cells, its enhancer activity strongly increased after 24 h of JQ1 treatment (Fig. 3e) and defined the single strongest gain in the entire region (Fig. 3f). We confirmed this effect in conventional reporter assays using a minimal *MYC* promoter (Fig. 3g) and found that it does not correlate with *PVT1* transcription (Extended Data Fig. 7b), indicating that this element primarily acts as a *MYC* enhancer. Collectively, our results suggest that the rebound of *MYC* is driven by a focal enhancer that is formed during acquired resistance and pre-established in primary resistant cells.

To search for regulatory pathways driving primary resistance, we compared steady-state transcriptomes of sensitive and resistant leukaemias. Among only 38 genes consistently upregulated in resistant leukaemias, we identified 17 genes previously implicated as Wnt signalling

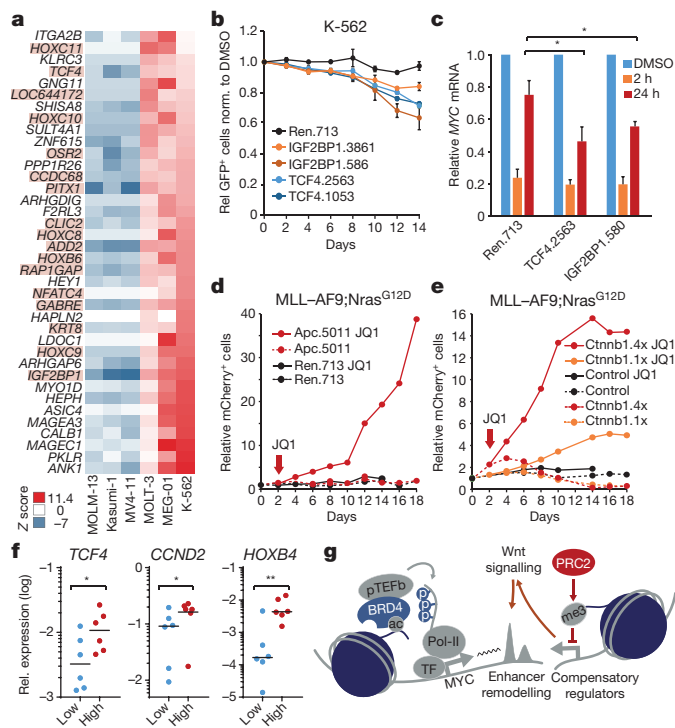


Figure 4 | Wnt signalling promotes primary and acquired BET resistance in leukaemia. **a**, Heat map of genes differentially expressed between indicated sensitive and resistant leukaemia cell lines. Genes previously implicated as Wnt target genes are highlighted in red. **b**, Competitive proliferation assay of K-562 cells expressing the indicated shRNAs. Shown is the fraction of GFP⁺/shRNA⁺ under JQ1 treatment (100 nM), normalized to the fraction in DMSO-treated control cells ($n = 3$, mean \pm s.e.m.). **c**, qRT-PCR analysis of *MYC* mRNA levels in K-562 cells expressing the indicated shRNAs at different time points after JQ1 treatment (200 nM) ($n = 3$; mean \pm s.e.m.; Student's *t*-test). **d**, Competitive proliferation assays of MLL-AF9;Nras^{G12D} leukaemia cells expressing the indicated shRNAs. Shown is the fraction of mCherry⁺/shRNA⁺ cells relative to the initial measurement. After 2 days, each sample was split in half, treated with DMSO or 50 nM JQ1 and followed up for 16 days. **e**, Competitive proliferation assays of MLL-AF9;Nras^{G12D} leukaemia cells co-expressing mCherry and Ctnnb1 harbouring one (Ctnnb1.1x) or four (Ctnnb1.4x) activating mutations. After 2 days, cells were treated with 50 nM JQ1 or DMSO, and the relative fraction of mCherry⁺/shRNA⁺ cells was followed over time. **f**, qRT-PCR analysis of *TCF4*, *CCND2* and *HOXB4* mRNA levels (relative to *GAPDH*) in sensitive ($n = 6$; $IC_{50} < 200$ nM) and resistant ($n = 6$; $IC_{50} > 500$ nM) primary human leukaemia samples (Student's *t*-test; * $P < 0.05$; ** $P < 0.01$). **g**, Model of transcriptional plasticity as a driver of primary and acquired BET resistance.

targets (Fig. 4a), which together with findings in our mouse model (Fig. 2f) pointed to Wnt signalling as a candidate driver of BET resistance. From these 17 genes, we selected two well-established Wnt targets, IGF2BP1²⁵ and TCF4²⁶, to test whether they contribute to the resistant state of K-562 cells. Suppression of TCF4 or IGF2BP1 increased the JQ1 sensitivity and diminished transcriptional rebound phenomena in K-562 cells (Fig. 4b, c and Extended Data Fig. 8a), while overexpression of TCF4 in MOLM-13 cells reduced their sensitivity (Extended Data Fig. 8b, c).

To investigate whether resistant cells engage Wnt transcriptional machinery following BET inhibition, we performed ChIP assays for TCF7L2, a Wnt-dependent transcription factor known to drive *MYC* and other Wnt targets in complex with CTNNB1²¹. Following JQ1 treatment, established Wnt target genes showed a marked increase in TCF7L2 binding, which was particularly prominent at the *PVT1* element 3' of *MYC* (Extended Data Fig. 8d), suggesting that this site acts as a Wnt-dependent *MYC* enhancer. To test whether Wnt activation drives *de novo* resistance, we transduced cells of our sensitive

AML mouse model with validated shRNAs targeting *Apc* (Extended Data Fig. 8e, f), a negative regulator of Wnt signalling, or active mutants of Ctnnb1. While Wnt activation had no or slightly detrimental effects in untreated cells, JQ1 treatment led to a rapid outgrowth of cells expressing two independent *Apc* shRNAs or Ctnnb1 containing four activating mutations (Ctnnb1.4x) (Fig. 4d, e and Extended Data Fig. 8g). Single-mutant Ctnnb1 had milder effects, indicating that the degree of BET resistance depends on Wnt activation levels. Ctnnb1.4x also completely blunted the response to JQ1 *in vivo*, and promoted resistance in two independent leukaemia models (Extended Data Fig. 8h, i).

After demonstrating that Wnt activation drives resistance in leukaemia, we wondered whether our Wnt expression signature would be more generally associated with BET resistance. Through integrating sensitivity profiles of 246 cell lines with available transcriptome data^{27,28}, we found that expression of this gene set is significantly increased in JQ1-resistant contexts (Extended Data Fig. 9a, b). As a first step towards probing Wnt as a BET resistance driver in other cancers, we found that suppression of *Apc* promotes resistance in a mouse model of pancreatic cancer, while treatment with the Wnt inhibitor pyrvinium²⁹ synergizes with JQ1 in this model and highly resistant ASPC-1 cells (Extended Data Fig. 9c–e). To probe whether Wnt signalling is associated with BET resistance in primary human leukaemia, we quantified nine Wnt-associated transcripts in sensitive and resistant patient-derived samples (Extended Data Fig. 10a). Notably, three of these transcripts (that is, *TCF4*, *CCND2* and *HOXB4*) were significantly overexpressed in resistant samples (Fig. 4f), while others showed a similar trend (Extended Data Fig. 10b). To reduce context-specific biases of single markers, we used the three significant transcripts to establish a simple 'resistance index', which strongly correlated with IC_{50} values (Extended Data Fig. 10c) and may provide a first step towards developing a predictive biomarker.

Through integrative profiling and functional genetic analyses in mouse and human leukaemia, our study reveals that BRD4 regulates a remarkably specific and diverse set of target genes. Leukaemia cells, through an adaptation process that is facilitated by inactivation of PRC2, can become resistant to BET inhibition by rewiring the transcriptional regulation of key BRD4 targets such as *MYC* (Fig. 4g). Interestingly, a recent study found that in T-lymphoblastic leukaemia *MYC* transcription can switch from a NOTCH1-dependent to a BRD4-dependent mode following treatment with γ -secretase inhibitors³⁰. Our study reveals that the BRD4-dependency of *MYC* can be overcome through engaging a focal enhancer and, together with another report³¹, establishes WNT signalling as a major driver and candidate biomarker of BET resistance in leukaemia. These findings highlight that the heterogeneity and plasticity of transcriptional machinery plays a major role in promoting resistance to chromatin-targeted therapies.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 10 November 2014; accepted 7 July 2015.

Published online 14 September 2015.

- Zuber, J. *et al.* RNAi screen identifies Brd4 as a therapeutic target in acute myeloid leukaemia. *Nature* **478**, 524–528 (2011).
- Dawson, M. A. *et al.* Inhibition of BET recruitment to chromatin as an effective treatment for MLL-fusion leukaemia. *Nature* **478**, 529–533 (2011).
- Delmore, J. E. *et al.* BET bromodomain inhibition as a therapeutic strategy to target c-Myc. *Cell* **146**, 904–917 (2011).
- Asangani, I. A. *et al.* Therapeutic targeting of BET bromodomain proteins in castration-resistant prostate cancer. *Nature* **510**, 278–282 (2014).
- Puissant, A. *et al.* Targeting MYCN in neuroblastoma by BET bromodomain inhibition. *Cancer Discov.* **3**, 308–323 (2013).
- Herait, P. E. *et al.* BET-bromodomain inhibitor OTX015 shows clinically meaningful activity at nontoxic doses: interim results of an ongoing phase I trial in hematologic malignancies. *Proc. 105th Annu. Meet. Am. Assoc. Cancer Res.* **Apr 5–9**, Abstract CT231 (2014).

7. Jang, M. K. *et al.* The bromodomain protein Brd4 is a positive regulatory component of P-TEFb and stimulates RNA polymerase II-dependent transcription. *Mol. Cell* **19**, 523–534 (2005).
8. Lovén, J. *et al.* Selective inhibition of tumor oncogenes by disruption of super-enhancers. *Cell* **153**, 320–334 (2013).
9. Roe, J.-S., Mercan, F., Rivera, K., Pappin, D. J. & Vakoc, C. R. BET bromodomain inhibition suppresses the function of hematopoietic transcription factors in acute myeloid leukemia. *Mol. Cell* **58**, 1028–1039 (2015).
10. Filippakopoulos, P. *et al.* Selective inhibition of BET bromodomains. *Nature* **468**, 1067–1073 (2010).
11. Fellmann, C. *et al.* An optimized microRNA backbone for effective single-copy RNAi. *Cell Rep.* **5**, 1704–1713 (2013).
12. De Raedt, T. *et al.* PRC2 loss amplifies Ras-driven transcription and confers sensitivity to BRD4-based therapies. *Nature* **514**, 247–251 (2014).
13. Neff, T. *et al.* Polycomb repressive complex 2 is required for MLL-AF9 leukemia. *Proc. Natl Acad. Sci. USA* **109**, 5028–5033 (2012).
14. Shi, J. *et al.* The Polycomb complex PRC2 supports aberrant self-renewal in a mouse model of MLL-AF9/Nras^{G12D} acute myeloid leukemia. *Oncogene* **32**, 930–938 (2013).
15. Zuber, J. *et al.* An integrated approach to dissecting oncogene addiction implicates a Myb-coordinated self-renewal program as essential for leukemia maintenance. *Genes Dev.* **25**, 1628–1640 (2011).
16. Di Croce, L. & Helin, K. Transcriptional regulation by Polycomb group proteins. *Nature Struct. Mol. Biol.* **20**, 1147–1155 (2013).
17. Rahman, S. *et al.* The Brd4 extraterminal domain confers transcription activation independent of pTEFb by recruiting multiple proteins, including NSD3. *Mol. Cell Biol.* **31**, 2641–2652 (2011).
18. Ballaré, C. *et al.* Phf19 links methylated Lys36 of histone H3 to regulation of Polycomb activity. *Nature Struct. Mol. Biol.* **19**, 1257–1265 (2012).
19. Shi, J. *et al.* Role of SWI/SNF in acute leukemia maintenance and enhancer-mediated Myc regulation. *Genes Dev.* **27**, 2648–2662 (2013).
20. Subramanian, A. *et al.* Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl Acad. Sci. USA* **102**, 15545–15550 (2005).
21. He, T. C. *et al.* Identification of c-MYC as a target of the APC pathway. *Science* **281**, 1509–1512 (1998).
22. Reya, T. *et al.* A role for Wnt signalling in self-renewal of haematopoietic stem cells. *Nature* **423**, 409–414 (2003).
23. Wang, Y. *et al.* The Wnt/ β -catenin pathway is required for the development of leukemia stem cells in AML. *Science* **327**, 1650–1653 (2010).
24. Arnold, C. D. *et al.* Genome-wide quantitative enhancer activity maps identified by STARR-seq. *Science* **339**, 1074–1077 (2013).
25. Noubissi, F. K. *et al.* CRD-BP mediates stabilization of β TrCP1 and c-myc mRNA in response to β -catenin signalling. *Nature* **441**, 898–901 (2006).
26. Kolligs, F. T. *et al.* ITF-2, a downstream target of the Wnt/TCF pathway, is activated in human cancers with beta-catenin defects and promotes neoplastic transformation. *Cancer Cell* **1**, 145–155 (2002).
27. Klijn, C. *et al.* A comprehensive transcriptional portrait of human cancer cell lines. *Nature Biotechnol.* **33**, 306–312 (2014).
28. Barretina, J. *et al.* The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. *Nature* **483**, 603–607 (2012).
29. Thorne, C. A. *et al.* Small-molecule inhibition of Wnt signaling through activation of casein kinase 1 α . *Nature Chem. Biol.* **6**, 829–836 (2010).
30. Knoechel, B. *et al.* An epigenetic mechanism of resistance to targeted therapy in T cell acute lymphoblastic leukemia. *Nature Genet.* **46**, 364–370 (2014).
31. Fong, C. Y. *et al.* BET inhibitor resistance emerges from leukaemia stem cells. *Nature* <http://dx.doi.org/10.1038/nature14888> (2015).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank M. Weißenböck, B. Hopfgartner and M. Fellner for technical support, S.-M. Kula and the IMP/IMBA Molecular Biology Service for help with library construction, G. Schmauß, T. Lendl, M. Weninger and G. Stengl and the Biooptics Service Facility for FACS; A. Sommer and his team at Campus Science Support Facilities (<http://www.csf.ac.at>) for Illumina sequencing, S. W. Lowe for reagents and discussions, and all members of the Zuber laboratory for reagents, protocols and discussions. This research was funded by a Starting Grant of the European Research Council (ERC no. 336860; to J.Z.), SFB grants F4704 and F4710 of the Austrian Science Fund (FWF), a Fellowship of the People Programme (Marie Curie Actions) of the European Union (to P.R.). F.M. is an EMBO long-term fellow (EMBO ALTF 491-2014) and research in the Stark laboratory is supported by an ERC Starting grant (no. 242922; to A.S.). The Zuber laboratory and research at the IMP is generously supported by Boehringer Ingelheim.

Author Contributions P.R. designed the shRNA library and performed the RNAi screen, P.R. and M.R. planned and performed most experiments and analysed data. T.N. and E.A. performed bioinformatics analysis of RNA- and ChIP-seq data. F.M., L.M.B. and A.S. performed and analysed STARR-seq experiments. M.M. generated validated BRD2/3/4 shRNAs and performed experiments. J.R.-S. and C.V. contributed ChIP-seq data and provided advice. B.P., S.C.-R. and P.V. performed analyses in primary patient samples. S.D. performed experiments in PDAC. T.H. improved protocols for shRNA deep-sequencing. J.J. contributed critical reagents and performed CFU assays. L.E.D. provided critical reagents. N.S., U.T.-G. and D.G. profiled JQ1 sensitivity and performed RNA-seq in human cell lines. M.P. and N.K. provided critical advice and support. J.Z. designed experiments, analysed data, and supervised this research. P.R., M.R. and J.Z. co-wrote the paper.

Author Information All RNA-seq, ChIP-seq and STARR-seq data have been deposited in the Gene Expression Omnibus under accession code GSE63782. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to J.Z. (zuber@imp.ac.at).

METHODS

Plasmids. The shRNA library used in the screen was cloned into pLMN (pMSCV-miR30-PGK-NeoR-IRES-GFP)¹⁵. The same vector was used for validation of primary hits and PRC2 core complex partners. Additional validation studies for Suz12 were performed using pRT3GEN-mir30 (pSIN-TRE3G-turboGFP-miR30-PGK-NeoR) or pLENC (pMSCV-miRE-PGK-NeoR-IRES-mCherry)¹¹. shRNAs targeting *Apc* were cloned into pLMPC (pMSCV-miRE-PGK-PuroR-IRES-mCherry). A *Ctnnb1* cDNA harbouring four activating mutations (S33A, S37A, T41A, S45A; *Ctnnb1.4x*) was obtained from Addgene (#24312) and cloned into pMSCV-IRES-mCherry. *Ctnnb1* harbouring a single mutation (S45P; *Ctnnb1.1x*) was cloned into pMSCV-IRES-GFP. shRNAs targeting *TCF4* and *IGF2BP1* were cloned into pRT3GEN (pSIN-TRE3G-turboGFP-miRE-PGK-NeoR)¹¹, and the *TCF4* cDNA was cloned into pMSCV-PGK-NeoR-IRES-GFP.

Antibodies. The following antibodies were used: Suz12 (39357, Active Motif), histone H3K27me3 (39155, Active Motif; 07-449, Millipore), histone H3K36me3 (Ab9050, Abcam), histone H3K27ac (ab4729, Abcam), histone H3 (61277, Active Motif), MYC (56058, Cell Signaling), *TCF4* (ab130014, Abcam), *IGF2BP1* (RN001M, MBL), *TCF7L2* (sc-8631, Santa Cruz), *BRD4* (A301-985A1002, Bethyl), *BRD3* (A302-583A, Bethyl), *BRD2* (A302-368A, Bethyl), *APC* (MABC202, Millipore) and β -actin (A3854, Sigma-Aldrich and ab49900, Abcam). Secondary antibodies were anti-mouse (926-32210, LI-COR), anti-rabbit (926-32211, LI-COR) and anti-goat (926-32214, LI-COR). Antibodies used for FACS were: PE anti-mouse CD117/c-Kit (Biolegend, 105808), APC anti-mouse CD11b/Mac-1 (101212, Biolegend), PE-Cy7 anti-mouse Ly-6G/Gr-1 (108416, Biolegend), PE-Cy7 anti-mouse Ly-6A/E/Sca-1 (108114, Biolegend), PerCP/Cy5.5 anti-mouse CD45.2 (109828, Biolegend), PE/Cy7 anti-mouse CD117/c-Kit (105814, Biolegend), APC anti-mouse CD150 (115925, Biolegend), PE-Cy7 anti-human CD11b/MAC-1 (101215, Biolegend) and APC anti-human CD36 (336207, Biolegend).

Pooled shRNA library screening. An shRNA library targeting 626 chromatin-associated mouse genes (Supplementary Information; shRNAs prim. screen performance) was assembled in pLMN¹⁵ by combining an existing chromatin-focused library (243 genes, 1094 shRNAs)¹ and 1793 additional shRNAs that were designed based on optimized sensor rules^{32,33}, and cloned, sequence-verified and pooled as previously described¹. After spiking in several control shRNAs at equimolar ratios, the final pool of 2917 pLMN-shRNAs was transduced in triplicate into the same MLL-AF9;Nras^{G12D}-driven model that we had previously used to identify *BRD4* as a target in AML¹. To ensure library representation a total of 60 million cells were infected with 5% transduction efficiency using conditions that predominantly lead to a single retroviral integration and represent each shRNA in a calculated number of > 800 cells. Throughout G418 drug selection (1 mg ml⁻¹) more than 50 million cells were maintained at each passage to preserve library representation. Two days after infection T0 samples were acquired (6 million GFP⁺ cells per replicate) using a FACS AriaII (BD Biosciences) and deep-sequencing analysis confirmed that the library was well represented in all replicates. After G418 drug selection for 7 days, T1 samples were obtained using FACS (6 million GFP⁺ cells per replicate), and cells were subsequently cultured in the presence of 0.5 mg ml⁻¹ G418 (15 million cells per replicate were maintained at each passage). To control for unspecific clonal events during the JQ1 resistance screen, stably selected cells were mixed in 4:1 ratio with G418-selected cells expressing an empty miR-30 cassette, mCherry and NeoR (pMSCV-miR30.empty-PGK-NeoR-IRES-mCherry). Replicates were subsequently treated for 4 days with 100 nM JQ1 followed by 50 nM JQ1 for 22 days. Culture medium was exchanged every 2 days until cells proliferated comparable to wild-type cells and afterwards passaged seven times. Three replicates cultured in the presence of vehicle (0.033% DMSO) were maintained in parallel. After 22 days (T2), about 6 million shRNA-expressing (GFP⁺) cells were sorted for each replicate by FACS. While mCherry⁺ cells completely disappeared over time, GFP⁺ cells survived and eventually grew in the presence of 50 nM JQ1 (corresponding to an IC₇₀ dose in this model), indicating that resistance did not emerge from random clonal events, but from shRNA-mediated effects. Genomic DNA from T1 and T2 samples was isolated by two rounds of phenol extraction using PhaseLock tubes (5PRIME), followed by isopropanol precipitation. Deep-sequencing libraries were generated by PCR amplification of shRNA guide strands using primers that tag the product with standard Illumina adapters (p7+Loop: CAAGCAGAAGACGGCATACGATAGTGAAGC CACAGATGT; p5+PGK: AATGATACGCGACACCGATGGATGTGGAAT GTGTGCGAGG). For each sample, DNA from at least 5 × 10⁶ cells was used as template in multiple parallel 50- μ l PCR reactions, each containing 0.5 μ g template, 1 × AmpliTaq Gold buffer, 0.2 mM of each dNTP, 2 mM MgCl₂, 0.3 μ M of each primer and 2.5 U AmpliTaq Gold (Life Technologies), which were run using the following cycling parameters: 95 °C for 10 min; 35 cycles of 95 °C for 30 s, 52 °C for 45 s and 72 °C for 60 s; 72 °C for 7 min. PCR products (340 bp) were combined for each sample, column purified using the QIAquick PCR purification kit (Qiagen) and further purified on a 1% agarose gel (QIAquick gel extraction kit, Qiagen). Libraries

were analysed on an Illumina HiSeq 2000 deep sequencer; 22 nucleotides of the guide strand were sequenced using a custom primer (miR30EcoRISeq, TAGCCCCCTT GAATTCGAGGAGCAGTAGGCA). To provide a sufficient baseline for detecting shRNA depletion in experimental samples, we aimed to acquire >500 reads per shRNA in the sequenced shRNA pool to compensate for variation in shRNA representation inherent in the pooled plasmid preparation or introduced by PCR biases. With these conditions, we acquired baselines of > 500 reads for all 2,917 shRNAs. Sequence processing was performed using a customized Galaxy platform³⁴. For each shRNA and condition, the number of matching reads was normalized to the total number of library-specific reads per lane and imported into a database for further analysis (Access 2013, Microsoft). All primary screen data are provided under Supplementary Information; shRNAs prim. screen performance.

Cell culture, retroviral gene transfer and RNAi studies. Mouse MLL-AF9;Nras^{G12D} AML cells (RN2)¹ and other murine leukaemia cell lines were derived from bone marrow of terminally diseased mice transplanted with fetal liver cells engineered to express the indicated oncogenes, as previously described³⁵. Murine leukaemia cells were cultured in RPMI 1640 (Gibco-Invitrogen) supplemented with 10% FBS, 20 mM glutamate, 10 mM sodium pyruvate, 10 mM HEPES (pH 7.3), 100 U ml⁻¹ penicillin and 100 μ g ml⁻¹ streptomycin. MV4-11, MEG-01 and K-562 cells were grown in IMDM with 10% FBS, 20 mM glutamate, 10 mM sodium pyruvate, 100 U ml⁻¹ penicillin and 100 μ g ml⁻¹ streptomycin. All human AML cell lines were cultured in RPMI 1640 (Gibco-Invitrogen) supplemented with 10% FBS, 20 mM glutamate, 10 mM sodium pyruvate, 100 U ml⁻¹ penicillin and 100 μ g ml⁻¹ streptomycin, except KASUMI-1 cells, which were cultured in 20% FBS. Human PDAC cell lines CAPAN-2 and MIAPACA-2 were cultured in DMEM (Gibco-Invitrogen) supplemented with 10% FBS, 20 mM glutamate, 10 mM sodium pyruvate, 100 U ml⁻¹ penicillin and 100 μ g ml⁻¹ streptomycin. SU-8686 and ASPC-1 were cultured in RPMI 1640 (Gibco-Invitrogen) supplemented with 10% FBS, 20 mM glutamate, 100 U ml⁻¹ penicillin and 100 μ g ml⁻¹ streptomycin. HUPT-4 and HPAF-2 were cultured in EMEM (Sigma-Aldrich) supplemented with 10% FBS, 20 mM glutamate, 100 U ml⁻¹ penicillin and 100 μ g ml⁻¹ streptomycin. Human SCLC cell lines DMS-273, SHP-77, NCI-H82 and NCI-H1048 were cultured in RPMI 1640 (Gibco-Invitrogen) supplemented with 10% FBS, 20 mM glutamate, 10% GlutaMAX, 100 U ml⁻¹ penicillin and 100 μ g ml⁻¹ streptomycin. Cell lines were obtained from ATCC (<http://www.lgcstandards-atcc.org/en.aspx>) or DSMZ (<http://www.dsmz.de/>) and tested for mycoplasma infection on a regular basis using a commercial biochemical test (Lonza).

Tet-on competent murine pancreatic cancer cells (KRPC2) were generated and characterized as previously described³⁶. In brief, pancreatic progenitor cells isolated from a murine fetus (ED17.5–18.5) harbouring a conditional endogenous Kras^{G12D} allele (lox-STOP-lox-Kras^{G12D})³⁷ and a conditional Trp53 deletion (*Trp53^{fl/fl}*)³⁸ were transduced with retroviral constructs expressing Myc, 4-OHT-inducible CreER^{T2}, codon-optimized Firefly Luciferase (Luc2), and rtTA3, and orthotopically injected into the pancreas of syngeneic recipient mice (detailed protocols available upon request). Emerging tumours were characterized to harbour histological features of human pancreatic adenocarcinoma and used to derive a cell line (KRPC2), which was cultured in DMEM (Gibco-Invitrogen) with 10% FBS, 20 mM glutamate, 10 mM sodium pyruvate, 100 U ml⁻¹ penicillin and 100 μ g ml⁻¹ streptomycin.

Retroviral packaging was performed using Platinum-E cells (Cell Biolabs) according to established protocols¹¹. In brief, for each calcium phosphate transfection, 10–20 μ g plasmid DNA and 5 μ g helper plasmid (pCMV-Gag-Pol, Cell Biolabs) were used. Transduction efficiencies of retroviral constructs were measured 48 h post infection by flow cytometry (Guava easyCyte, Millipore). Transduced cell populations were usually selected 48 h post infection using 0.5 mg ml⁻¹ G418 (Gibco, Life Technologies). To generate resistant MLL-AF9;Nras^{G12D} murine leukaemia, cells were transduced with retrovirally delivered shRNAs or cDNAs at a transduction efficacy between 5–15% (predominantly yielding single viral integrations per cell). JQ1 treatment was started post infection and medium was carefully exchanged every 2 days to maintain a 50 nM JQ1 concentration until cells were proliferating normally and daily passaging was required. The starting point of JQ1 treatment is critical and can vary between different targets and shRNAs since the effect is dependent on shRNA potency, protein half-life of the respective target and the half-life of the associated chromatin mark. In case of strong detrimental effects the optimal start of JQ1 treatment needs to be determined for every target.

MOLM-13 and K-562 cells were modified to express the ecotropic receptor and rtTA3 using retroviral transduction of pMSCV-RIEP (pMSCV-rtTA3-IRES-EcoR-PGK-Puro) or lentiviral transduction of pWPXLd-RIEP (pWPXLd-rtTA3-IRES-EcoR-PGK-Puro) followed by puromycin selection (0.5 and 1 μ g ml⁻¹, respectively, for 1 week). Derived cell lines were subsequently transduced with ecotopically packaged retroviruses.

For flow cytometry immunophenotyping, cultured cells were collected and stained using the indicated antibodies (two million cells in 100 μ l FACS buffer:

PBS, 5% FCS, 0.005% sodium acetate, and human or mouse TruStain fX Fc-receptor blocking agent (Biolegend) for 20 min at 4 °C. Antibodies were diluted 1:400. Stained samples were analysed on an LSR Fortessa (BD) flow cytometer. Data analysis was performed using FlowJo software (Treestar). Cell viability assays for drug IC₅₀ and GI₅₀ determination were performed using Alamar blue staining according to manufacturer's guidelines, or by counting the increase in viable cell numbers over 72 h in the presence of different JQ1 concentrations. Dead cells were excluded using propidium iodide (PI) staining. Measurements of cell concentration were performed on a Guava easyCyte (Millipore), gating only viable cells (FSC/SSC/PI⁻). Synergy was evaluated using Chou-Talalay Combination Index (CI), calculated with the CompuSyn program (ComboSyn, Inc.)³⁹. GI₅₀ values for all tested cell lines are provided in Supplementary Information; JQ1 profiled cell lines.

Competitive proliferation assays using shRNAs in pLMN, pLENC, pLMP or pRT3GEN-mir30/RT3GEN vectors were performed as described previously^{1,40}. Briefly, for assays involving Tet-regulated shRNA expression vectors, Tet-on competent MLL-AF9;Nras^{G12D} mouse AML cells⁴⁰ or Tet-on competent derivatives of human leukaemia cell lines were transduced with the respective plasmid at infection rates of <5% to ensure single-copy integration. For constitutive shRNA expression (pLMN) cells were analysed for GFP⁺ expression 1 day post infection. pRT3GEN transduced cells were selected for 7 days using G418 (0.5 mg ml⁻¹), mixed with 10% uninfected cells, and shRNA expression was induced through addition of doxycycline (DOX) to a final concentration of 1 µg ml⁻¹. The percentage of shRNA-expressing cells (turboGFP⁺ or GFP⁺) was measured daily using flow cytometry. All values were normalized to day 1. Once the percentage of viable cells was below 20%, measurements were discontinued (indicated in the graph by the discontinuation of the respective sample). Sequences for all shRNAs used in single assays are provided in Supplementary Information; shRNAs single assays.

Colony-forming assays were performed in Methocult (StemCell Technologies, Cat. No. M3231). 1000 AML cells (in 100 µl standard medium) were added to 900 µl Methocult supplemented with 1 µl of DMSO or JQ1 (to a final concentration of 200 nM). After 7 days the types of colonies were enumerated and normalized to input.

Gene expression and protein level analysis. RNA was prepared using the RNeasy Plus Mini Kit (Qiagen). Synthesis of cDNA was performed using SuperScript III Reverse Transcriptase (Invitrogen) or Taqman reverse transcription kit (Applied Biosystems) with random hexamers. Quantitative PCR analysis was performed on an ABI 7900HT with SYBR green (ABI). All signals were quantified using the $\Delta\Delta C_T$ method and were normalized to the levels of β -actin or GAPDH. All primers used in this study are presented in Supplementary Information; Primer.

For immunoblotting of Suz12, histone H3K27me3, BRD4, TCF4, IGF2BP1 and MYC, 20 µg of whole-cell lysate (lysis buffer: 50 mM TRIS pH 8, 250 mM NaCl, 0.5% NP-40, 5 mM EDTA) were loaded onto each lane. Protein extracts were resolved by SDS polyacrylamide gel electrophoresis (SDS-PAGE) and transferred to nitrocellulose for blotting.

Preparation of RNA-seq libraries. RNA from human cell lines or mouse AML cells was isolated using the RNeasy Plus Mini Kit (Qiagen). Messenger RNA was obtained using two rounds of poly(A) selection using the Dynabeads mRNA purification kit (Invitrogen) and subsequently fragmented by incubation at 94 °C for 3 min (fragmentation buffer: 40 mM Tris base adjusted to pH 8.2 with glacial acetic acid, 100 mM potassium acetate, 30 mM magnesium acetate in DEPC-treated H₂O). Cell lines infected with shRNA expressing vectors were FAC-sorted for shRNA expression (GFP⁺) before RNA extraction. The fragmented mRNA was used as template for first-strand cDNA synthesis with random hexamers and a NEBNext DNA Library Prep Master Mix Set for Illumina Synthesis kit (Invitrogen). The second-strand cDNA was synthesized with 100 mM dATP, dCTP, dGTP and dUTP in the presence of RNase H, *E. coli* DNA polymerase I and DNA ligase (Invitrogen). The incorporation of dUTP allowed elimination of the second strand during library preparation (described below) and thus preservation of strand specificity.

Chromatin immunoprecipitation and ChIP-seq library construction. ChIP assays were performed exactly as described¹⁴ in two to three independent biological replicates. Crosslinking was performed with sequential EGS and formaldehyde or with formaldehyde alone. All samples were quantified by quantitative PCR performed using SYBR green (ABI) on an ABI 7900HT after crosslink reversal. Each immunoprecipitate signal was referenced to an input standard-curve dilution series (immunoprecipitate/input) to normalize for differences in starting cell number and for primer amplification efficiency and normalized to a control region. Two independent biological replicates were performed for each ChIP-seq experiment.

For ChIP-Seq library construction, 1×10^7 leukaemia cells were crosslinked using 1% formaldehyde for 20 min at room temperature. After purifying immunoprecipitated DNA using QIAquick Gel Extraction Kit (QIAGEN), ChIP-seq lib-

raries were constructed using the TruSeq ChIP Sample Prep Kit (Illumina) following manufacturer's instructions with the following exception: following adaptor ligation, libraries were amplified for 15 cycles. In brief, purified ChIP DNA was repaired to blunt ends followed by dA-tailing process to ligate adapters for multiplex-based sequencing. Size-selection of adaptor-ligated ChIP-DNA was performed with agarose-gel-based selection of DNA size ranged from 200 to 350 bp. The final libraries were amplified for 15 cycles, and size-selected with SPRI clean-up by using AMPure XP beads (Beckman Coulter). Library quality was determined using a Bioanalyzer (Agilent). Libraries sizes ranged from 250 to 300 bp. ChIP-seq libraries were sequenced using an Illumina HiSeq 2000 platform. Barcoded libraries were sequenced in a multiplexed fashion with two to six libraries at equal molar ratio, with single-end reads of 50 bases.

Illumina deep sequencing. Two to five nanograms of cDNA or DNA precipitated by ChIP was used as starting material for the generation of single-end sequencing libraries as described in Illumina's ChIP-seq sample preparation protocol. DNA fragments of the following sizes were selected for these experiments: 200–350 bp for ChIP-seq, 150–700 bp for RNA-seq. For strand-specific RNA-seq, the uridine residues present in one cDNA strand were digested with uracil-N-glycosylase (New England Biolabs), followed by PCR amplification. Completed libraries were quantified on a Bioanalyzer using the dsDNA 1000 assay kit (Agilent) and the qPCR NGS library quantification kit (Agilent). Cluster generation and sequencing was performed using a HiSeq 2000 system with a read length of 50 or 100 nucleotides according to the manufacturer's guidelines (Illumina).

Analysis of RNA-seq data. The calculation of RNA expression values was based on the RefSeq database, which was downloaded from UCSC on 10 January 2014. Genes with overlapping exons were flagged and double entries (that is, exactly the same gene at two different genomic locations) were renamed. Identical genes with more than one assigned gene symbol were flagged. Genes with several transcripts were merged to consensus genes consisting of a union of all underlying exons using the fuge software (I. Tamir, unpublished), which resulted in 25,098 gene models. Gene names and accession numbers of identical genes are presented in Supplementary Information; refseq.hg19.2014_0110.imp.map.

Paired-end and single-read fragments were trimmed on their 5' end (six cycles PE100, two cycles SR50 NEB RNA sample prep protocol), adaptors were removed using cutadapt v1.4.2⁴¹ and reads with a length of less than 18 bp in any read of the pair were discarded. The trimmed and adaptor-free reads were aligned against rDNA of the respective organism using bowtie2 v2.0.2⁴² for paired-end and bowtie v0.12.5⁴³ for single-end reads. The rRNA cleaned paired-end reads were aligned against the transcriptome to estimate the fragment size and the standard deviation of the fragment size using bwa v0.6.2⁴⁴. The rRNA cleaned reads were aligned to the genome with the TopHat splice junction mapper for RNA-seq reads⁴⁵. The uniquely aligning reads were used for counting per gene with htseq-count v0.6.1p⁴⁶ with the overlap-resolution mode option set to 'union' using the processed RefSeq-annotated gene database. Reads per kilobase per million mapped reads (RPKM) values for all samples were calculated as in ref. 47. Read counts for all samples were normalized using the median-of-ratios method implemented in DESeq2, v1.2.10⁴⁸. Counts were transformed for PCA using DESeq2's variance stabilizing transformation. For both the resistant and sensitive groups, three independent RNA-seq experiments on different leukaemia cell lines were available. The edgeR R package v3.4.2⁴⁹ was used to calculate the significance of difference in expression of resistant and sensitive leukaemia cell lines (DMSO treated) in Fig. 4a. For this, raw counts were normalized using the 'relative log expression' method analogous to DESeq2. Dispersions were sequentially estimated with the default settings by first estimating a common dispersion for the gene set and then 'shrinking' tag-wise dispersion values towards that dispersion trend. Statistical modelling and analysis was done by fitting a model taking the separation into sensitive and resistant cell lines into account. Differential expression of genes between those two groups was subjected to statistical analysis and resulting *P* values were adjusted for multiple testing with the Benjamini-Hochberg correction⁵⁰. The resulting list was further filtered by only considering genes which all have higher normalized counts in the sensitive cell lines versus the resistant cell lines or vice versa.

Gene set enrichment analysis. Gene set enrichment analysis²⁰ was performed using GSEA v2.07 software with 1,000 gene set permutations. KEGG signalling gene sets and Wnt signalling associated gene sets were obtained from the Molecular Signatures Database v4.0 (MSigDB, <http://www.broadinstitute.org/gsea/msigdb/index.jsp>). To perform GSEA with human gene signatures, human-mouse homologues from the mouse gene sets were identified and converted into human gene names using the NCBI homologue database (build68). Ambiguous homologue mapping, that is, different mouse gene names pointing to the same human homologue, was excluded. A detailed description of GSEA methodology and interpretation is provided at <http://www.broadinstitute.org/gsea/doc/GSEAUUserGuideFrame.html>. In brief, the normalized enrichment score (NES) provides "the degree to which a gene set is overrepresented at the top or bottom

of a ranked list of genes". The nominal P value (Nom p -val) describes "the statistical significance of the enrichment score". The false discovery rate q value (FDR q -val) is "the estimated probability that a gene set with a given NES represents a false positive finding". In general, given the lack of coherence in most expression data sets and the relatively small number of gene sets being analysed, an FDR cut-off of 25% is appropriate. All gene sets used in this study are provided in Supplementary Information; GSEA gene sets.

ChIP-seq data analysis. Adaptors of the single-read fragments were removed using cutadapt v1.4.2⁴¹ and reads with a length of less than 18bp in any read of the pair were discarded. The remaining reads were aligned to the genome (mouse: mm10, human: hg19) using Bowtie v1.0.0⁴³. Reads with same start and end position on the same strand were removed from the alignment. To identify ChIP-seq peaks, we used the MACS2 v2.1.0.20141030 (<https://github.com/taoliu/MACS/>) peak finding algorithm. A threefold enrichment relative to input control samples was used for peak calling as well as the option to call broad peaks. Building a shifting model was disabled and the small nearby and large nearby region parameters were set to 5,000 and 20,000 respectively. The extension size was set to the respective median insert size of the ChIP-seq treatment sample for paired-end data and the estimated fragment size for single-end data. Read numbers for the resulting peaks were quantified using the BEDTools suite⁵¹ and normalized to total mapped reads. The Brd4 ChIP-seq track was obtained from published data sets¹⁹ (GEO sample GSM1262345). Peaks were assigned to modified versions of the RefSeq database (GRCh37/hg19 downloaded on 10 January 2014 from UCSC Table Browser and GRCm38/mm10 downloaded on 30 June 2014 from UCSC Table Browser) by assigning peaks to the closest upstream/downstream TSS using BEDTools. For these versions, overlapping genes were flagged and double entries (that is, exactly the same gene at two different genomic locations) were renamed. Identical genes with more than one assigned gene symbol were flagged. For the density heat maps, ChIP-seq peaks were identified using the MACS version 1.4.0beta (Model based Analysis of ChIP-Seq) peak finding algorithm⁵². A P value threshold of enrichment of 1×10^{-5} , a false discovery rate (FDR) of less than 1%, and a fivefold enrichment relative to input control samples were used for all peak calling. The Brd4 peaks were considered as promoter if the peak showed at least 1 bp overlap within ± 200 bp of RefSeq gene TSSs. If Brd4-called peaks showed no overlap within ± 200 bp of RefSeq gene TSSs, they were considered as enhancer-bound peaks. Heat map matrices were created by counting tags using a 10 kb window (± 5 kb of the peak summit) and 20 bp bin size. Further visualization of heat map matrices were done using Java TreeView 1.1.6r4 (<http://jtreeview.sourceforge.net>). For comparing DMSO and JQ1 treated ChIP-seq data sets, heat map matrices were normalized to total mapped reads counts before visualization with Java TreeView.

STARR-seq screens. STARR-seq screens were done as described previously²⁴, with the following exceptions. DNA from 21 bacterial artificial chromosomes (BACs), which were available for the extended *MYC* locus covering approximately 91% of the surrounding 3 Mb of the genomic region, plus 25 genic control BACs, were used for library generation. The fragmented BAC DNA was cloned into the human STARR-seq screening vector²⁴ containing a minimal *MYC* promoter. We transfected 500 μ g of this library into 1×10^8 K-562 cells per condition, using a BTX Agile Pulse MAX system (Harvard Apparatus) following the manufacturer's protocol. Electroporation conditions: two pulses at 750 V for 0.5 ms (100 ms interval), followed by 5 pulses of 250 V for 2 ms (100 ms interval) in BTXpress cytoporation medium T. After electroporation, the cells were grown at 37 °C in RPMI-1640 containing either DMSO or 250 nM JQ1 and were harvested after 24 h. For each screen we isolated poly(A) RNA and processed it as described before²⁴, with a different extension time of 45 s in the final amplification of the cDNA and modified primers for the first cDNA amplification. All BACs used for the STARR-seq screen are provided in Supplementary Information; STARR-seq BACs.

STARR-seq deep sequencing and analysis. All sequencing was performed as a 50-cycle paired-end run on an Illumina HiSeq2000 machine. STARR-seq and input read processing was performed as described²⁴, with the following exceptions. Libraries from different conditions were normalized to each other by subsampling to 10^5 reads per condition using reservoir sampling. Peaks were called using macs2 (v2.1.0)⁵² with the input library as background. The peaks were then filtered for a minimal fold enrichment of three and only peaks with at least five reads per kb were considered for further analysis. Peak calling was performed for STARR-seq reads from both conditions (DMSO and JQ1) and the resulting peaks were merged using bedtools merge⁵¹. STARR-seq enrichment was calculated using bedtools coverage⁵¹. Differential fold enrichment was calculated as the ratio between enrichments in JQ1 over DMSO and expressed in \log_2 .

Luciferase assays and data analysis. Luciferase assays were done as described previously²⁴, with the following exceptions. Enhancer candidate regions were amplified from genomic DNA and shuttled into a modified pGL4 luciferase reporter vector containing a minimal *MYC* promoter (chr8:128,748,440-

128,748,550). Each construct was tested by co-transfecting 10^6 K-562 cells with 45 μ g of the firefly luciferase construct and 5 μ g of a renilla luciferase control plasmid (pGL4.75, Promega). After electroporation, the cells were grown at 37 °C in RPMI-1640 containing either DMSO or 250 nM JQ1 and were harvested after 46 h. Relative luciferase signals (firefly/renilla) were normalized to signal from two negative background sequences chosen based on their absence of STARR-seq signal and expressed as fold change over background. We analysed the statistical significance of differential luciferase signals with or without JQ1 treatment using an unpaired t -test with Welch's correction for three biological replicates. Primer sequences which were used to amplify the particular enhancer regions are deposited in Supplementary Information; Primer.

Animal experiments. Mouse MLL-AF9;Nras^{G12D} leukaemia cells (RN2) were transplanted by tail-vein injection of 1×10^5 cells into sub-lethally (5.5 Gy) irradiated female B6/SJL (CD45.1) recipient mice at the age of 6–10 weeks ($n = 5$, for each individual experiment). For whole-body bioluminescent imaging, mice were intraperitoneally injected with 50 mg kg⁻¹ D-luciferin (Goldbio), and after 5 min were analysed using an IVIS Spectrum system (Caliper LifeSciences). For JQ1 treatment trials, a stock of 100 mg ml⁻¹ JQ1 in DMSO was 20-fold diluted by dropwise addition of a 10% 2-hydroxypropyl- β -cyclodextrin carrier (Sigma) under vortexing, yielding a 5 mg ml⁻¹ final solution. Mice were intraperitoneally injected daily with freshly diluted JQ1 (50 mg kg⁻¹) or a similar volume of carrier containing 5% DMSO. JQ1 treatment was started 3 days after transplantation of RN2 cells stably transduced with pLMN shRen.713 and shSuz12.1676. Mice transplanted with RN2 cells expressing β -catenin cDNA or empty vector were treated with JQ1 daily starting 1 day after transplantation. All animals were maintained in the pathogen-free animal facility of the Research Institute of Molecular Pathology, and all procedures were carried out according to an ethical animal license, which is approved and regularly controlled by the Austrian Veterinary Authorities. No statistical methods were used to predetermine sample size. The investigators were not blinded to allocation during experiments and outcome assessment. Randomization was not applied because all animals used in this study were similar for age, sex and strain background.

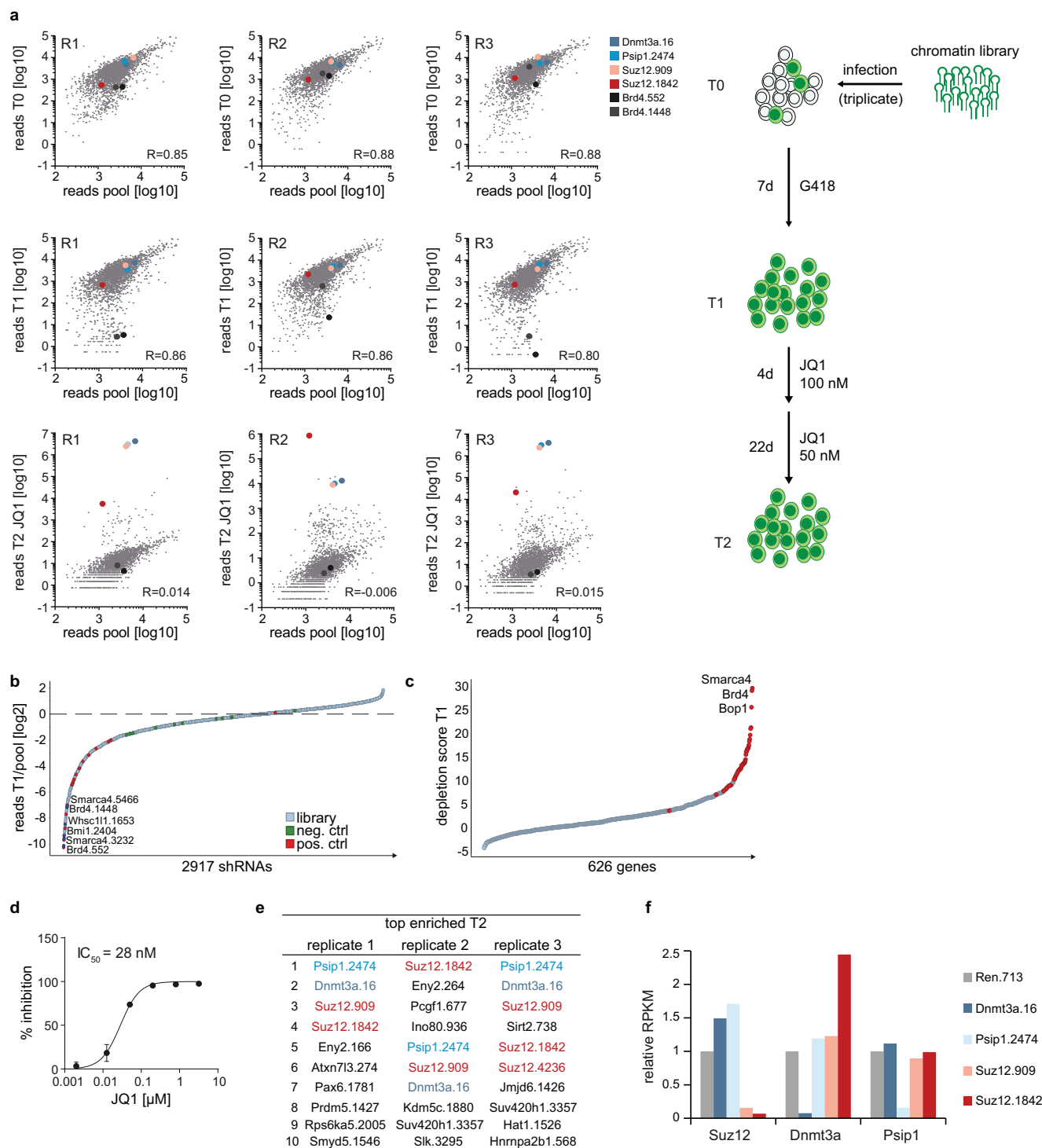
Patient sample analysis. Ten patients with AML (females, $n = 4$; males, $n = 6$) and two with CML blast phase (BP) (males, $n = 2$) were examined. Diagnoses were established according to the proposal of the French-American-British (FAB) cooperative study group^{53,54} and the classification of the World Health Organization (WHO)⁵⁵. Informed consent was obtained before bone marrow puncture in each case. The study was approved by the Institutional Review Board (Ethics Committee) of the Medical University of Vienna. Mononuclear cells (MNC) were isolated using Ficoll and stored in liquid nitrogen until used. After thawing, the viability of cells ranged from 75% to 85% as assessed by trypan blue exclusion. RNA was isolated from bone marrow or peripheral blood MNC using the RNeasy MinElute CleanupKit (Qiagen). cDNA was synthesized using RQ1 DNase buffer, RQ1 DNase and RQ1 DNase Stop Solution (Promega), Moloney murine leukaemia virus reverse transcriptase, random primers, First Strand buffer, dNTPs (100 mM), and RNasin (Invitrogen) according to the manufacturer's instructions. Details of the patient material used is deposited in Supplementary Information; Patient information.

MNC were cultured in 96-well plates (TPP) in RPMI-1640 medium (Lonza) with 10% fetal calf serum in the absence or presence of JQ1 (50–5,000 nM) at 37 °C for 48 h. AML MNC were incubated with JQ1 in the presence of 100 ng ml⁻¹ recombinant human (rh) G-CSF (Amgen), 100 ng ml⁻¹ rhSCF (Peprotech) and 100 ng ml⁻¹ rhIL-3 (Novartis Pharma AG). CML BP cells were incubated with JQ1 without cytokines. After incubation, 0.5 μ Ci 3H-thymidine (Perkin Elmer) was added for 16 h. Cells were then harvested on filter membranes in a Filtermate 196 harvester (Packard Bioscience). Filters were air-dried and the bound radioactivity was measured in a Top-Count NXT β -counter (Packard Bioscience). All experiments were performed in triplicates. Proliferation was calculated as a percentage of medium control, and the inhibitory effects of JQ1 were expressed as IC₅₀ values.

Statistical analyses. Results are expressed as means \pm standard error of the mean. If not stated otherwise statistical significance was calculated by two-tailed unpaired t -test on two experimental conditions with $P < 0.05$ considered statistically significant. Statistical significance levels are denoted as follows: **** $P < 0.0001$; *** $P < 0.001$; ** $P < 0.01$; * $P < 0.05$. No statistical methods were used to predetermine sample size.

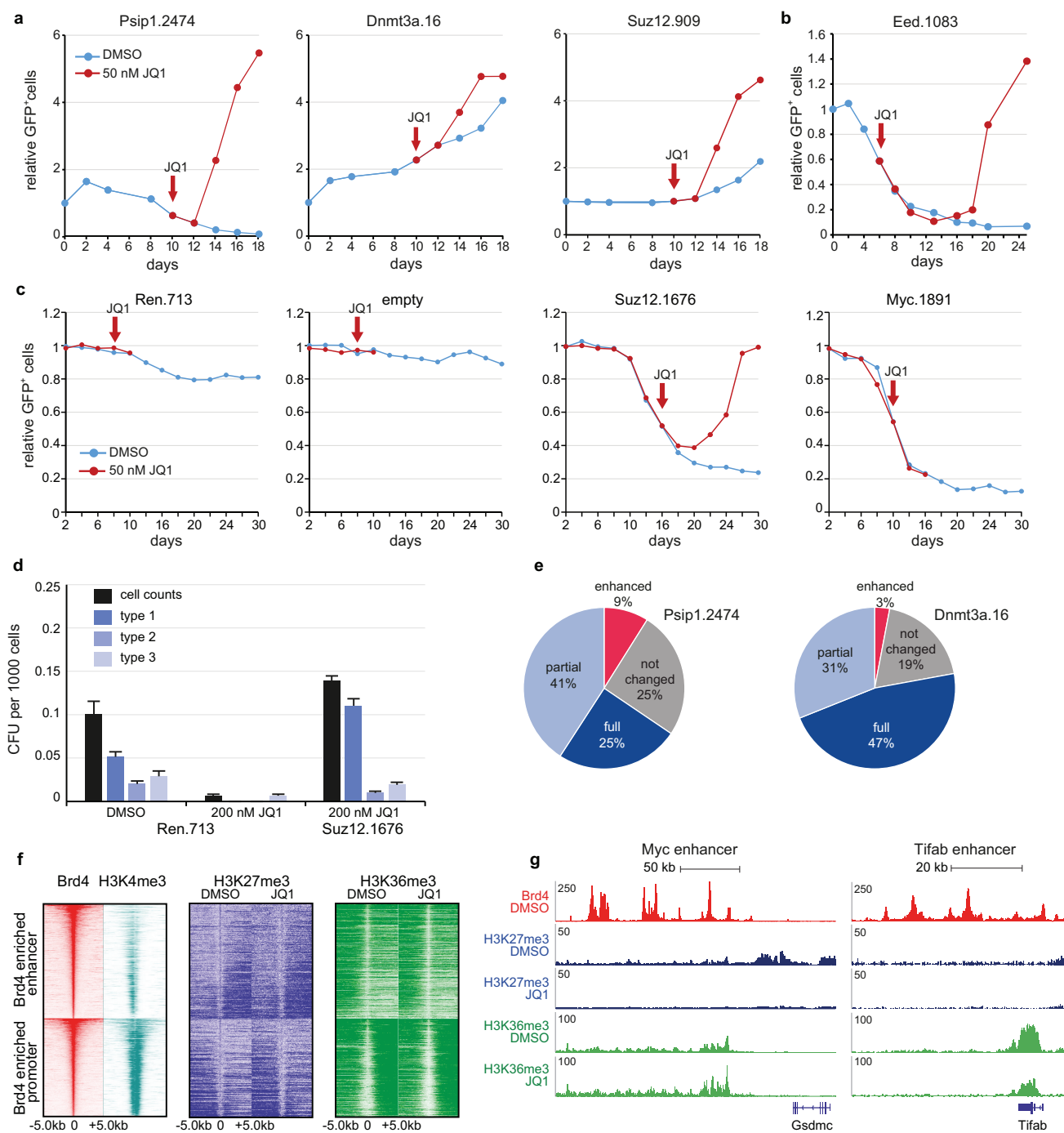
32. Fellmann, C. *et al.* Functional identification of optimized RNAi triggers using a massively parallel sensor assay. *Mol. Cell* **41**, 733–746 (2011).
33. Dow, L. E. *et al.* A pipeline for the generation of shRNA transgenic mice. *Nature Protocols* **7**, 374–393 (2012).
34. Goecks, J., Nekutenko, A. & Taylor, J. Galaxy: a comprehensive approach for supporting accessible, reproducible, and transparent computational research in the life sciences. *Genome Biol.* **11**, R86 (2010).

35. Zuber, J. *et al.* Mouse models of human AML accurately predict chemotherapy response. *Genes Dev.* **23**, 877–889 (2009).
36. Lito, P. *et al.* Disruption of CRAF-mediated MEK activation is required for effective MEK inhibition in KRAS mutant tumors. *Cancer Cell* **25**, 697–710 (2014).
37. Jackson, E. L. *et al.* Analysis of lung tumor initiation and progression using conditional expression of oncogenic K-ras. *Genes Dev.* **15**, 3243–3248 (2001).
38. Jonkers, J. *et al.* Synergistic tumor suppressor activity of BRCA2 and p53 in a conditional mouse model for breast cancer. *Nature Genet.* **29**, 418–425 (2001).
39. Chou, T. C. Theoretical basis, experimental design, and computerized simulation of synergism and antagonism in drug combination studies. *Pharmacol. Rev.* **58**, 621–681 (2006).
40. Zuber, J. *et al.* Toolkit for evaluating genes required for proliferation and survival using tetracycline-regulated RNAi. *Nature Biotechnol.* **29**, 79–83 (2011).
41. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *Embnnet/journal* **17**, 10 (2011).
42. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nature Methods* **9**, 357–359 (2012).
43. Langmead, B., Trapnell, C., Pop, M. & Salzberg, S. L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* **10**, R25 (2009).
44. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).
45. Trapnell, C., Pachter, L. & Salzberg, S. L. TopHat: discovering splice junctions with RNA-seq. *Bioinformatics* **25**, 1105–1111 (2009).
46. Anders, S., Pyl, P. T. & Huber, W. HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics* **31**, 166–169 (2015).
47. Mortazavi, A., Williams, B. A., McCue, K., Schaeffer, L. & Wold, B. Mapping and quantifying mammalian transcriptomes by RNA-seq. *Nature Methods* **5**, 621–628 (2008).
48. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).
49. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139–140 (2010).
50. Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. B* **57**, 289–300 (1995).
51. Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842 (2010).
52. Zhang, Y. *et al.* Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* **9**, R137 (2008).
53. Bennett, J. M. *et al.* Proposed revised criteria for the classification of acute myeloid leukemia. A report of the French–American–British cooperative group. *Ann. Intern. Med.* **103**, 620–625 (1985).
54. Bennett, J. M. *et al.* Proposals for the classification of the acute leukaemias. French–American–British (FAB) co-operative group. *Br. J. Haematol.* **33**, 451–458 (1976).
55. Vardiman, J. W. *et al.* The 2008 revision of the World Health Organization (WHO) classification of myeloid neoplasms and acute leukemia: rationale and important changes. *Blood* **114**, 937–951 (2009).



Extended Data Figure 1 | Multiplexed shRNA screening for chromatin-associated dependencies in AML maintenance and BET resistance. **a**, Scatter plot illustrating the correlation of normalized reads per shRNA at all three time points (T0, T1 and T2; top to bottom) compared to the plasmid pool in all three independent replicates (left to right). The schematic to the right illustrates the different sampling time points subjected to deep sequencing. Top hits enriched under JQ1 treatment and positive controls from the initial negative-selection screen (T1) are marked with coloured dots according to the legend on the right. **b**, Pooled negative-selection screening depicting changes in representation of 2,917 shRNAs during 7 days of culture. shRNA fold depletion values were calculated by dividing the number of reads after 7 days of culture (T1) by the number of reads obtained from the plasmid pool, and are plotted as the mean of three replicates in ascending order. Completely depleted shRNAs

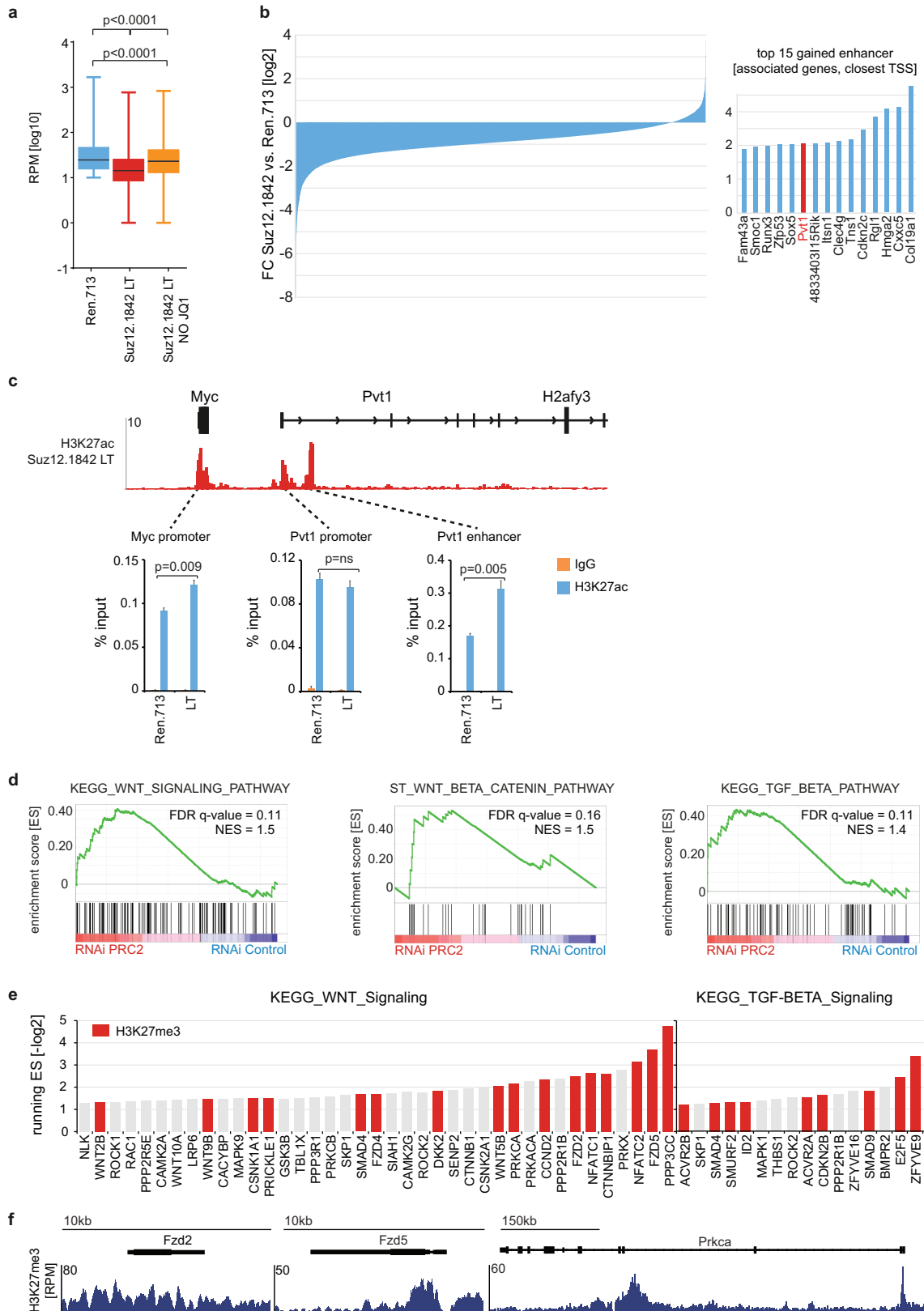
(0 reads at T1) obtained a fold depletion value of 1×10^{-3} . Positive control shRNAs targeting essential genes are marked in red; negative control shRNAs are depicted in green. **c**, Scatter plot depicting all genes ranked by the sum of their average depletion score of all shRNAs across all three replicates. Top scoring hits were defined as genes for which at least two shRNAs showed an average depletion of eightfold after 7 days of shRNA expression and are marked in red (45 genes). **d**, IC_{50} determination for JQ1 in murine MLL-AF9;Nras^{G12D} AML cells (RN2). Obtained numbers of viable cells per ml were normalized to DMSO ($n = 3$, mean \pm s.e.m.). **e**, Table showing the top ten enriched shRNAs at T2. shRNAs targeting Suz12, Dnmt3a and Psip1 are strongly enriched in all three independent replicates. **f**, Relative mRNA abundance (RPKM) of shRNA target genes in JQ1-resistant leukaemia cells expressing the indicated shRNAs, plotted relative to leukaemia cells expressing Ren.713.



Extended Data Figure 2 | PRC2 suppression confers resistance to JQ1.

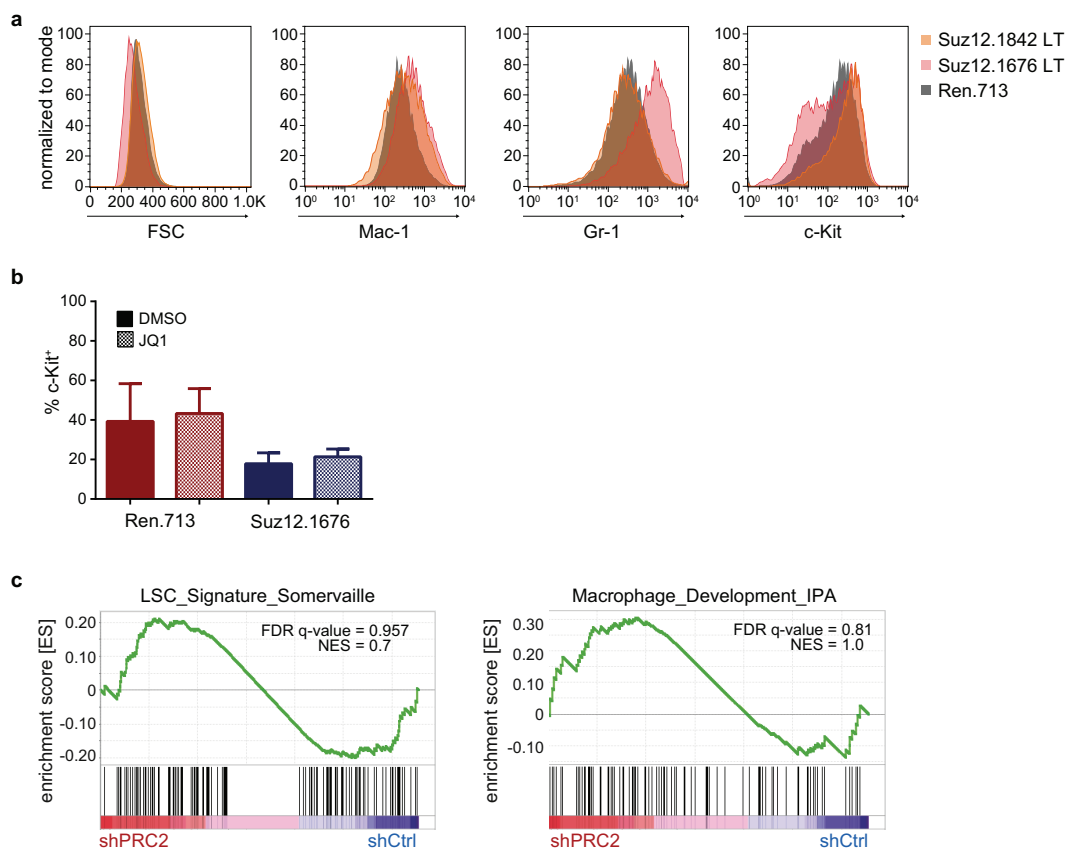
a, Competitive proliferation assays of MLL–AF9;Nras^{G12D} leukaemia cells transduced with pLMN constructs expressing the indicated shRNAs. Shown is the fraction of GFP⁺/shRNA⁺ cells (relative to the initial ratio) under treatment with JQ1 (50 nM) or DMSO over time (continuation from Fig. 1d). JQ1 treatment was initiated at the indicated time points (red arrow). **b**, Competitive proliferation assays evaluating one validated shRNA targeting Eed as an additional core component of the PRC2 complex (as in **a**). **c**, Competitive proliferation assays of Tet-on competent MLL–AF9;Nras^{G12D} leukaemia cells expressing the indicated shRNAs from a Tet-inducible vector (pRT3GEN-miR30). Transduced cells were selected with G418 and subsequently mixed with wild-type (wt) cells in a ratio of 95% to 5%. shRNA-expression was induced using dox treatment (1 µg ml⁻¹; from day 0), and the fraction of GFP⁺ cells was measured over time and plotted relative to day 2. JQ1 treatment (50 nM) was initiated in one of the duplicate samples at the indicated time point (red arrow). Once the percentage of viable cells was below 10%, measurements were discontinued (indicated in the graph by the discontinuation of the respective sample). The negative effect induced by Suz12 suppression is reverted upon treatment with JQ1, which is not the case when Myc is suppressed. **d**, Bar chart showing colony-forming cell

frequencies of MLL–AF9;Nras^{G12D} leukaemia cells expressing Ren.713 or Suz12.1676 shRNAs in the presence of DMSO or 200 nM JQ1; type 1, myeloblasts; type 2, maturing myeloblasts; type 3, terminally differentiated myeloid cells ($n = 3$, mean \pm s.e.m.). **e**, Pie charts depicting the fraction of JQ1-response genes which are re-expressed to the indicated extent in mouse resistant AML cells expressing shRNAs targeting Dnmt3a and Psip1 (continuation of Fig. 2b). JQ1-response genes (Fig. 2a) were grouped into four categories based on the divergence of their expression in resistant AML compared to AML expressing Ren.713 (not changed compared to expression after 24 h JQ1, less than 1.5-fold; restoration relative to DMSO control: full restoration, less than 1.5-fold; enhanced, above 1.5-fold; partial restoration, restored but less than 1.5 fold) **f**, Heat map showing the distribution of H3K27me3, H3K36me3 and H3K4me3 ChIP-seq peaks (fold enrichment >5 over input, FDR <1%) relative to Brd4 enhancer and promoter binding sites in MLL–AF9;Nras^{G12D} AML cells with and without JQ1 treatment. **g**, ChIP-seq occupancy profiles of Brd4 and H3K27me3 and H3K36me3 chromatin marks at enhancer regions downstream of *Myc* and upstream of *Tifab* following 3 days of treatment with vehicle or JQ1 (25 nM). y axis reflects the number of normalized cumulative tag counts in each region.



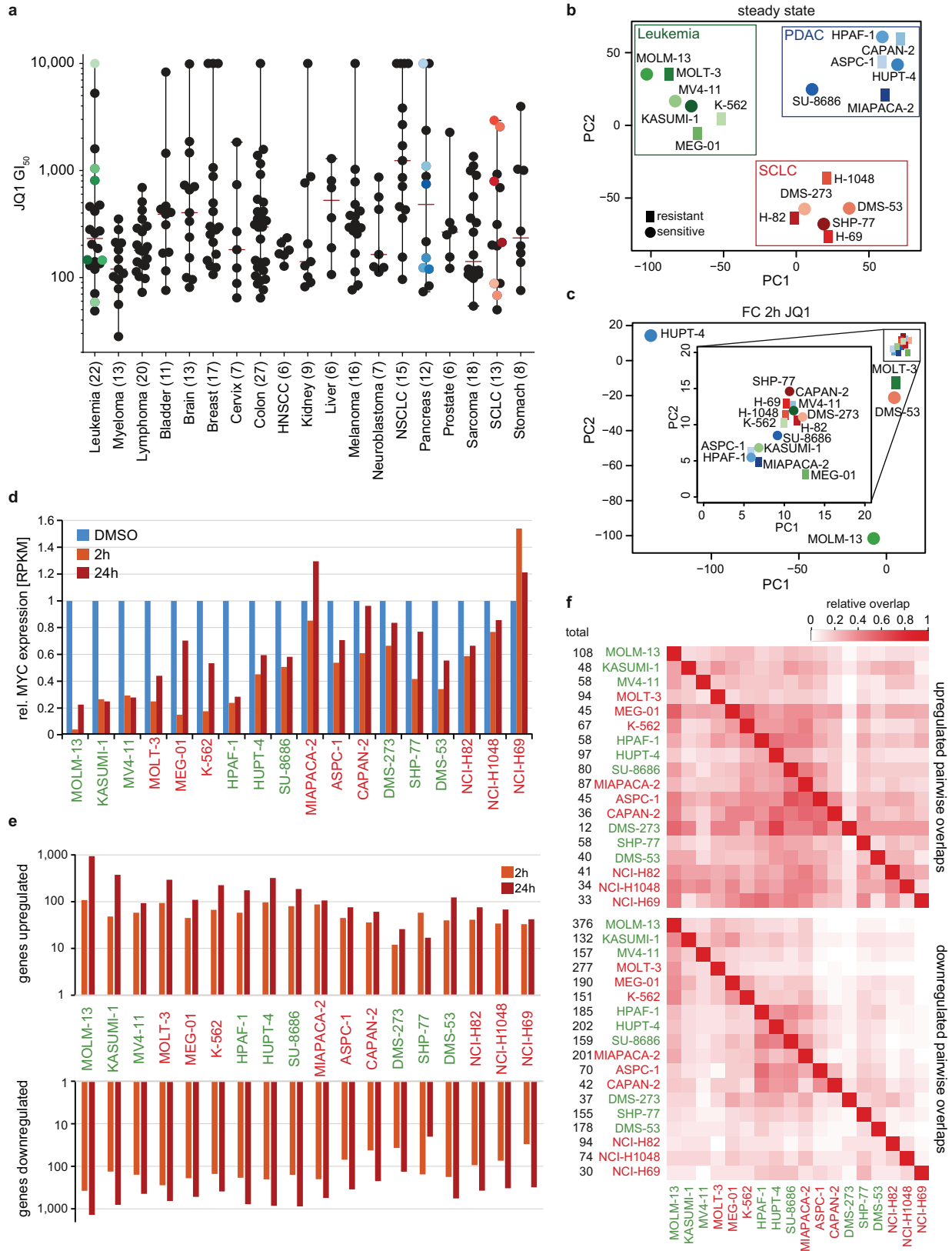
Extended Data Figure 3 | Changes in the regulatory landscape of BET-inhibitor-resistant mouse AML cells. **a**, Global H3K27 acetylation density of Suz12.1842-expressing resistant MLL–AF9;Nras^{G12D} leukaemia cells under long term (LT) treatment with 50 nM JQ1 (red bar), after 4 days of drug withdrawal (orange bar) and in Ren.713 controls (blue bar; statistical significance determined using Student's *t*-test). **b**, Left panel, sorted fold change (FC) ratios of H3K27ac peaks in long-term JQ1-treated MLL–AF9;Nras^{G12D} leukaemia cells expressing Suz12.1842 compared to cells expressing Ren.713 control shRNA. Included were all peaks showing > 10 reads per million in at least one condition. Right panel, top 15 gained peaks and their associated genes (defined using the closest transcription start site, TSS). The *Myc* proximal enhancer in the first intron of *Pvt1* is highlighted in red as one of the most differentially enriched peaks (FC=4.18). **c**, qRT–PCR validation of presented

H3K27ac ChIP-seq at the indicated regions downstream of the *Myc* locus ($n = 3$, mean \pm s.e.m., statistical significance determined using Student's *t*-test). **d**, Gene set enrichment analysis plots of three publicly available gene sets associated with signalling pathways comparing expression changes in resistant MLL–AF9;Nras^{G12D} AML cells induced by suppression of PRC2 complex members, compared to control cells expressing Ren.713 shRNA ($n = 2$) or empty vector (continuation of Fig. 2e, f). **e**, Core signature genes of KEGG-curated Wnt and TGF- β gene sets with increased expression in resistant murine MLL–AF9;Nras^{G12D} cells, compared to sensitive cells. Red coloured bars indicate association with H3K27 methylation in JQ1-sensitive MLL–AF9;Nras^{G12D} AML cells. **f**, H3K27 methylation density at three exemplified genes with high expression changes in JQ1-resistant murine AML.



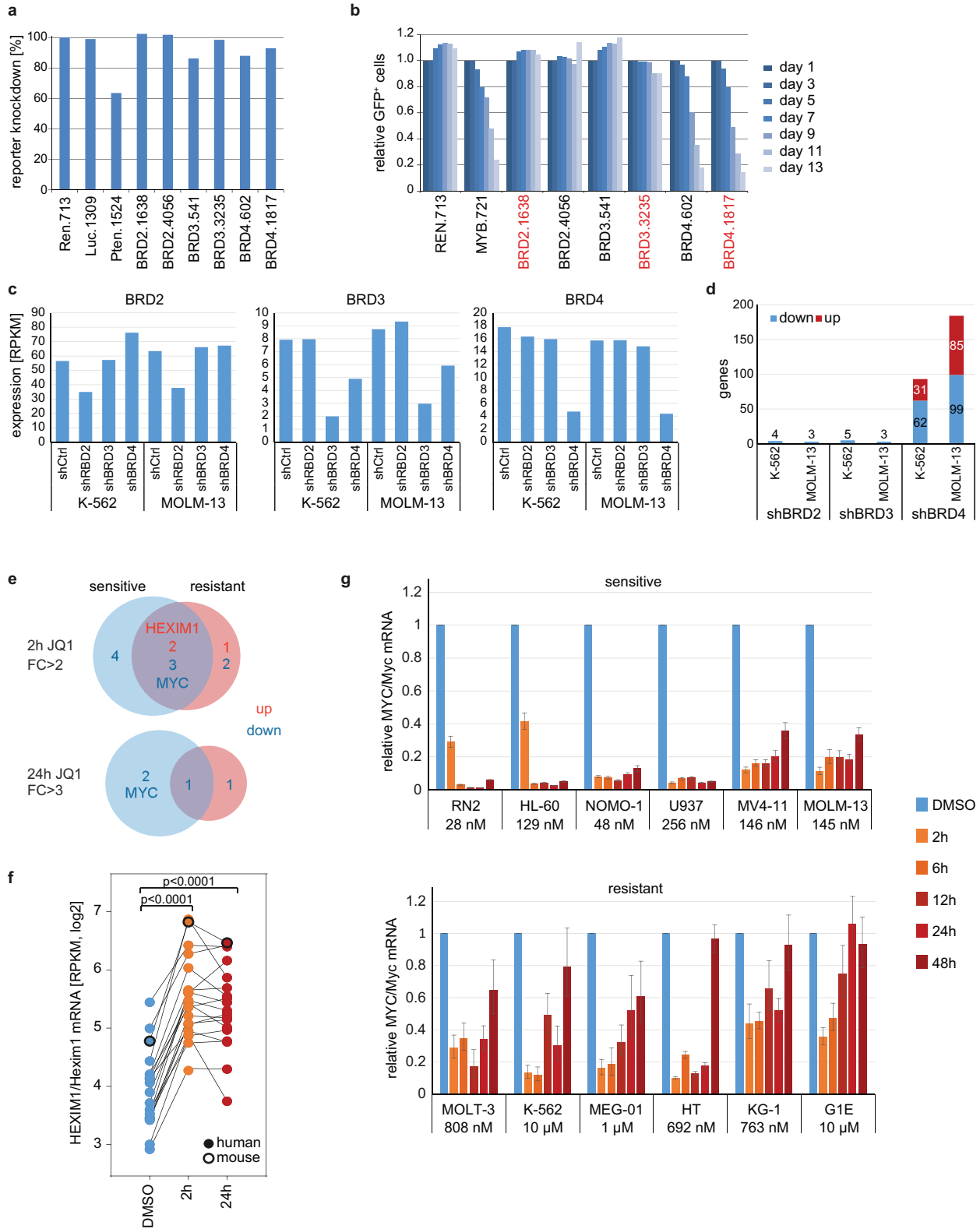
Extended Data Figure 4 | Resistant MLL-AF9;Nras^{G12D} AML cells generated through Suz12 suppression are not enriched for LSC-associated surface markers or expression signatures. **a**, Immunophenotyping of JQ1-resistant mouse AML cells expressing two independent Suz12 shRNAs stably cultured in 50 nM JQ1 for more than 4 weeks (LT) compared to cells expressing Ren.713 control shRNA. Data are representative of two independent biological replicates. **b**, Percentage of c-Kit⁺ cells in CD45.2⁺ bone marrow

cells isolated from terminally diseased CD45.1⁺ mice following transplantation with MLL-AF9;Nras^{G12D} cells expressing Ren.713 or Suz12.1676 and *in vivo* treatment with DMSO carrier or JQ1 (50 mg kg⁻¹ per day) ($n = 5$, mean \pm s.e.m.). **c**, Gene set enrichment analysis evaluating changes in macrophage and LSC gene signatures in resistant MLL-AF9;Nras^{G12D} AML expressing three PRC2 shRNAs compared to cells expressing Ren.713 shRNA ($n = 2$) or empty vector (see also Fig. 2e, f).



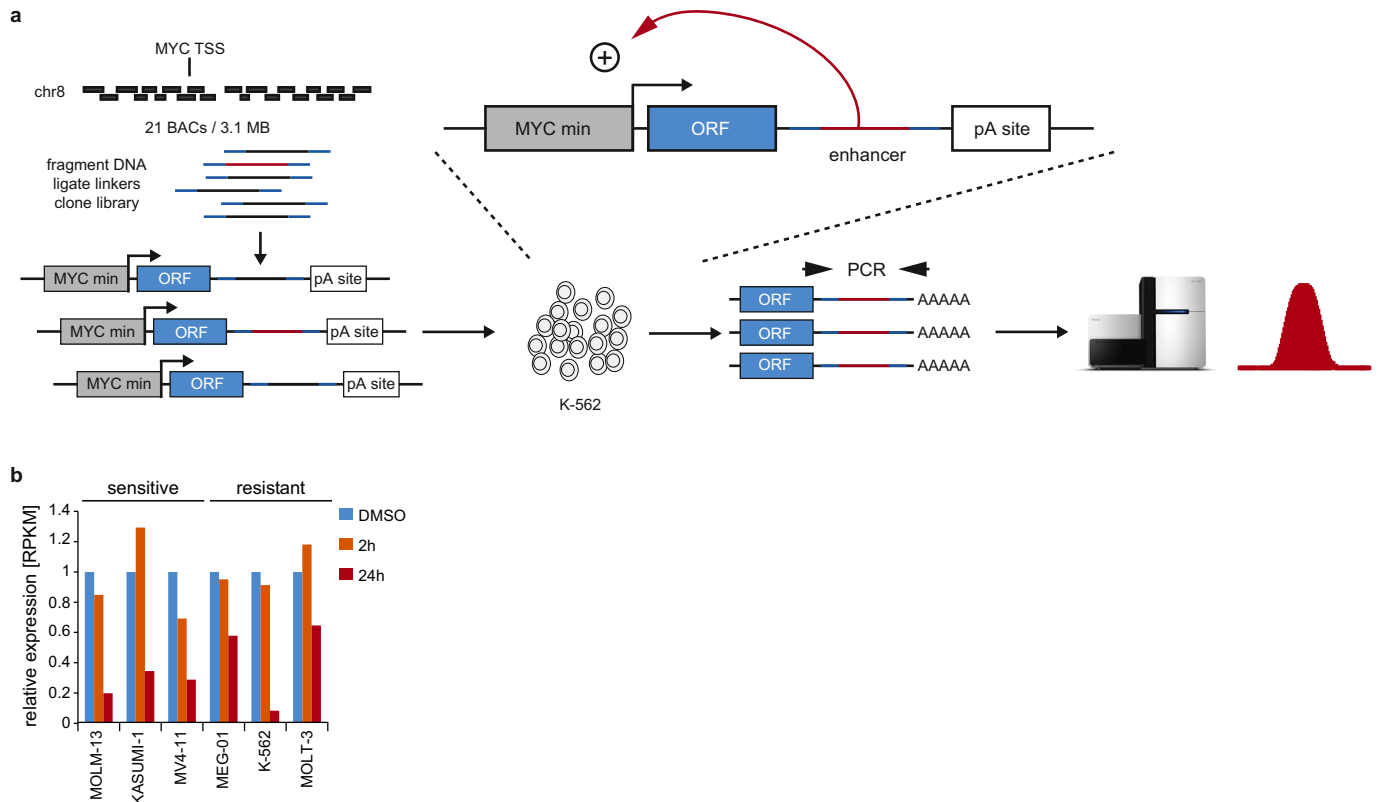
Extended Data Figure 5 | Comparison of JQ1-response genes in sensitive and resistant cancer cell lines. **a**, JQ1 sensitivity profiling in 246 human cancer cell lines of different tissue contexts. Shown are GI_{50} values determined using Alamar blue staining after 72 h. **b**, Principal component analysis of steady-state transcriptomes (based on RPKM) and **c**, transcription changes (based on fold change, FC) following 2 h of JQ1 treatment (200 nM) in indicated sensitive and resistant cancer cell lines of different tissue context. Steady-state profiles cluster based on tissue context, while neither baseline nor dynamic expression analysis can accurately distinguish sensitive and resistant contexts. **d**, *MYC* mRNA levels (RPKM) at different time points after JQ1 treatment (200 nM) in indicated cell lines, relative to levels in DMSO-treated

cells. Individual cell line pairs are grouped for their tissue context and coloured according to their sensitivities (green, sensitive; red, resistant). **e**, Number of genes twofold up- or downregulated upon JQ1 treatment (200 nM) after 2 h and 24 h (minimum expression >3 RPKM). **f**, Pairwise overlap of commonly up- or downregulated genes after 2 h of JQ1 treatment (200 nM) relative to DMSO control. Each cell colour corresponds to the relative number of commonly up- or downregulated genes in the cell lines listed in the respective row and column. The total number of genes regulated per cell line is indicated in black next to each cell line. Only little overlap is observed between cell lines of the same tissue context as well as between JQ1-sensitive or -resistant cell lines.



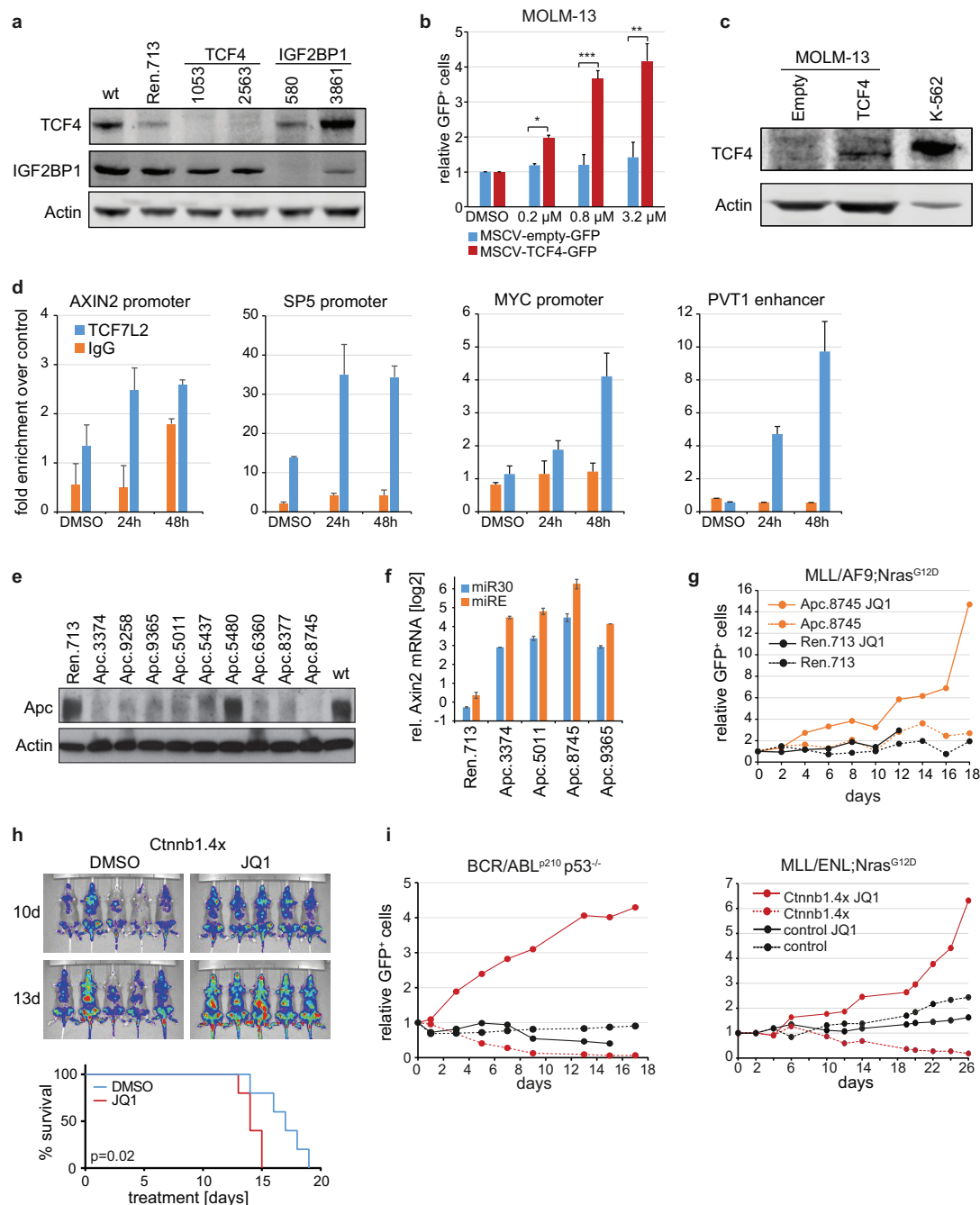
Extended Data Figure 6 | shRNA-based analysis of BRD2/3/4-dependent target genes and detailed analysis of effects on *HEXIM1* and *MYC* transcription in different cell lines. **a**, Determination of knockdown levels of a set of BRD2, BRD3 and BRD4 shRNAs determined using an established fluorescence-based shRNA reporter assay¹¹. Knockdown levels were quantified relative to Ren.713 control shRNA. **b**, Competitive proliferation assays in MOLM-13 cells functionally evaluating the potency of BRD2, BRD3 and BRD4 shRNAs over time using a Tet-regulated vector (pRT3GEN) in presence of doxycycline. Only shRNAs targeting BRD4 induce a proliferative disadvantage and lead to rapid depletion of GFP⁺ cells over time. Red labels indicate most potent shRNAs based on results obtained from reporter and competitive proliferation assay. **c**, Determination of *BRD2*, *BRD3* and *BRD4* mRNA levels in K-562 and MOLM-13 cells expressing the indicated shRNA. **d**, Number of genes commonly up- or downregulated with a fold change (FC) >3 in MOLM-13 or K-562 cells following expression of indicated

validated shRNAs or treatment with 200 nM JQ1 for 24 h. JQ1-induced expression changes show the largest overlap with cells expressing a validated BRD4 shRNA, suggesting that suppression of BRD4 is the key effector mechanism of JQ1 in leukaemia. **e**, Venn diagrams showing the overlap of genes commonly up- or downregulated following 2 h and 24 h of JQ1 either in all contexts, or specifically in sensitive or resistant leukaemias. **f**, *HEXIM1/Hexim1* expression (RPKM) in all 18 analysed human cell lines and murine MLL-AF9;Nras^{G12D} AML cells after 2 h and 24 h treatment with 200 nM JQ1, compared to DMSO control (statistical significance was determined using a paired Student's *t*-test). **g**, Relative *MYC* mRNA levels determined by qRT-PCR quantification in the indicated cell lines after incubation with 200 nM JQ1 measured at different treatment time points. Cell lines are grouped according to their sensitivity and the respective IC₅₀ is presented below. Resistant cell lines rapidly restore *MYC* transcription ($n = 3$, mean \pm s.e.m.).



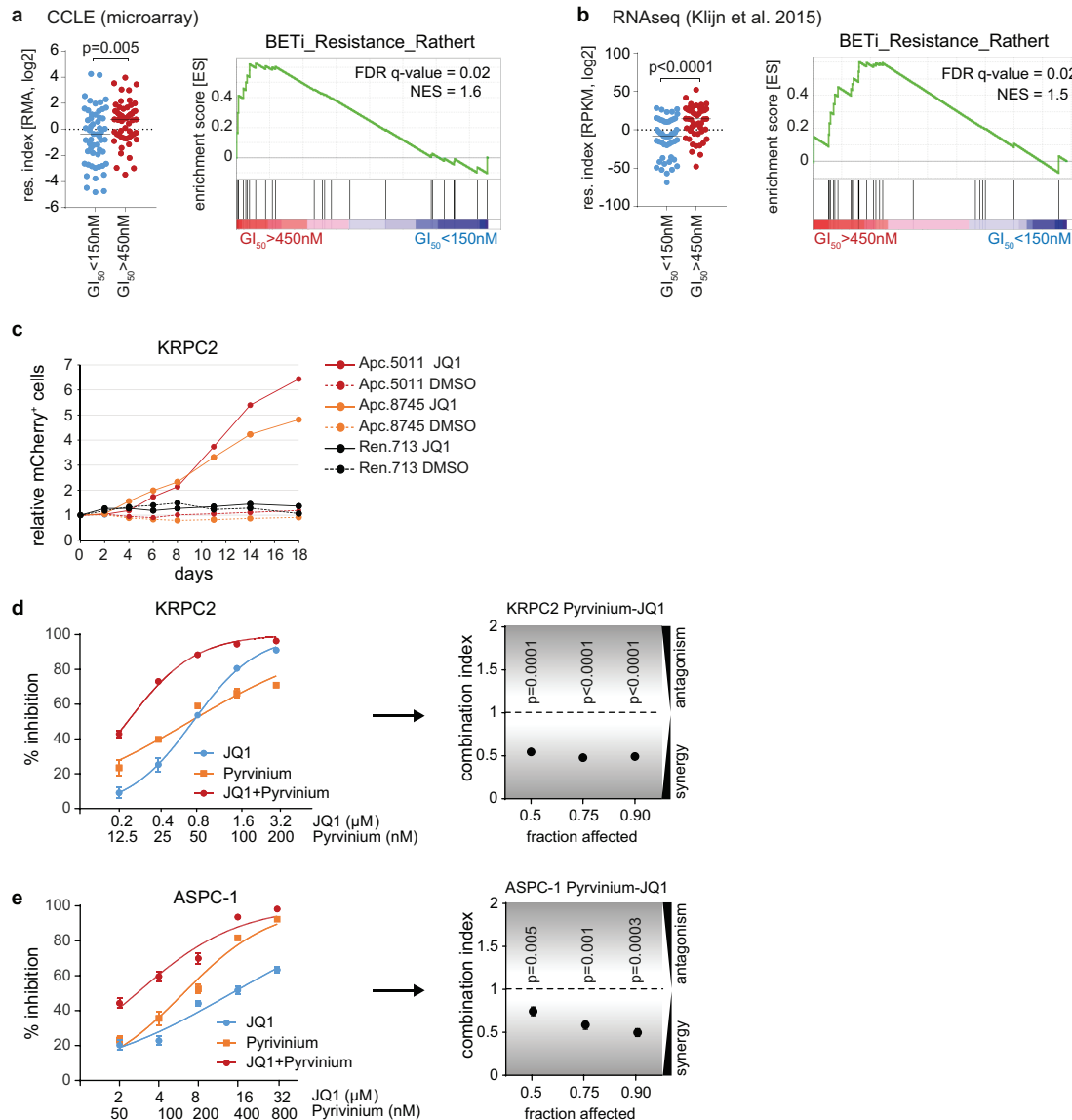
Extended Data Figure 7 | Dynamic STARR-seq analysis of enhancer activity in K-562 cells. **a**, Schematic representation of the STARR-seq cloning and screening strategy in K-562 cells. BACs available for the extended *MYC* locus (covering approximately 91% of a 3.1 Mb region at the *MYC* locus) and 25 genic control BACs were fragmented and cloned into a modified STARR-seq vector containing a minimal *MYC* promoter. This library was then screened for enhancer activity using STARR-seq in K-562 cells with or without 250 nM JQ1. The schematic shows the underlying principle of STARR-seq: a

bona fide enhancer can activate its own transcription from a minimal *MYC* promoter. Messenger RNA corresponding to active enhancer elements will therefore become more abundant among the cellular RNA compared to inactive fragments. **b**, *PVT1* mRNA levels (RPKM) at different time points after JQ1 treatment (200 nM) in indicated cell lines, relative to levels in DMSO-treated cells. *PVT1* expression is generally reduced upon JQ1 treatment indicating no association with enhancer activation.



Extended Data Figure 8 | Analysis and functional validation of Wnt signalling as a key driver of BET resistance. **a**, Protein levels of TCF4 and IGF2BP1 in K-562 cells transduced with pRT3GEN expressing indicated shRNAs after 7 days of doxycycline treatment, compared to Ren.713 and wild-type (wt) control samples. **b**, Competitive proliferation assay of JQ1-sensitive MOLM-13 cells expressing GFP-linked TCF4 cDNA or empty vector. Plotted is the relative fraction of GFP⁺ cells 72 h after JQ1 treatment using the indicated doses ($n = 3$, means \pm s.e.m., *** $P < 0.001$; ** $P < 0.01$; * $P < 0.05$ as determined by Student's t -test). Cells overexpressing TCF4 exhibit a dose-dependent competitive advantage under JQ1 treatment. **c**, Protein levels of TCF4 after overexpression of TCF4 cDNA subcloned into pMSCV-PGK-NeoR-IRES-GFP in MOLM-13 after 4 weeks of G418 (0.5 mg ml^{-1}) selection, compared to MOLM-13 transduced with empty control vector and to TCF4 protein levels in K-562 cells. **d**, ChIP qRT-PCR analysis of TCF7L2 binding to *AXIN2*, *SP5*, *MYC* promoter and the *PVT1* enhancer element in K-562 cells at indicated time points after treatment with 200 nM JQ1 ($n = 2$ biological replicates, mean \pm s.e.m.). TCF7L2 binding increases gradually over time at promoters of Wnt target genes and the *PVT1* enhancer at the *MYC* locus.

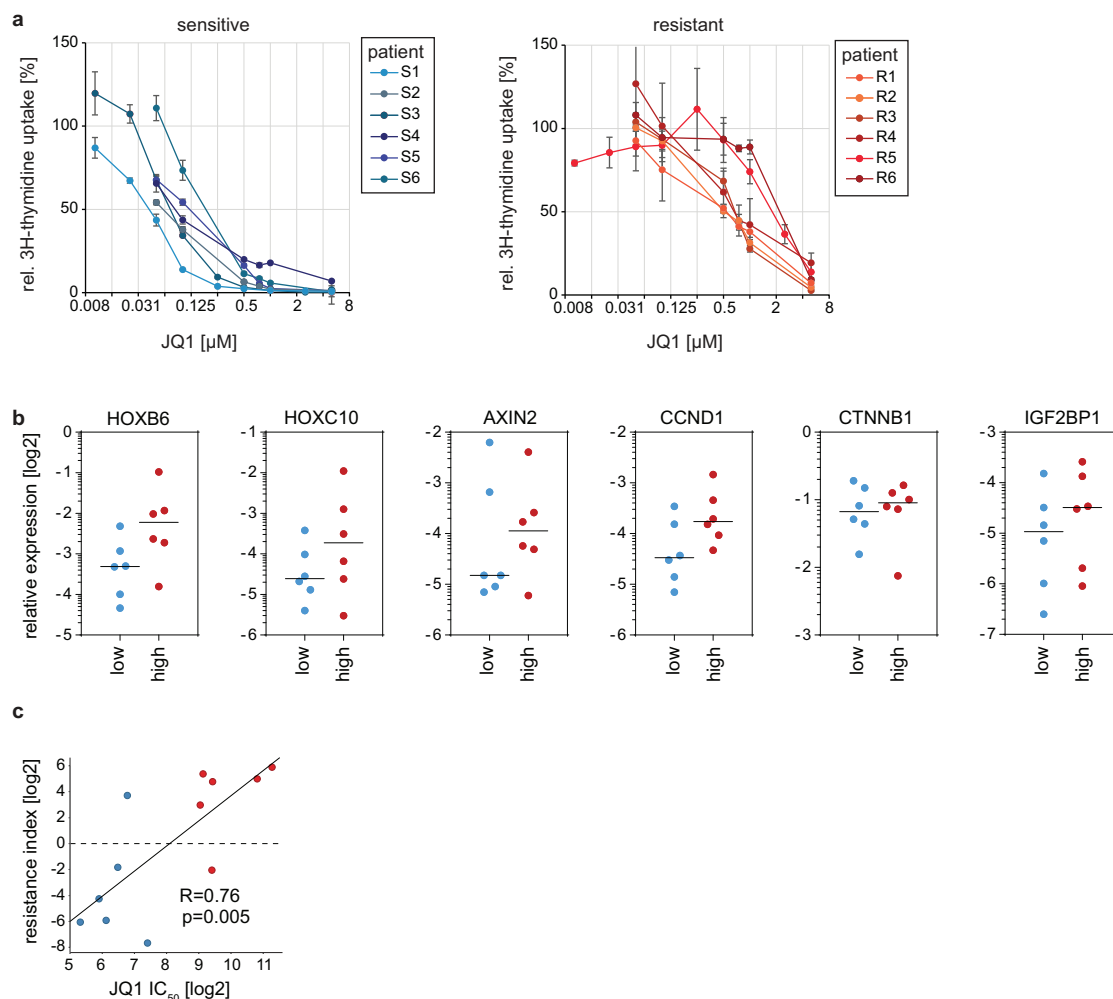
e, Protein levels of Apc in 3T3 murine fibroblast cells 7 days after infection with the indicated shRNAs cloned into pLMP compared to Ren.713 and wild-type control samples. **f**, Relative *Axin2* mRNA expression levels, determined by qRT-PCR normalized to B2m, after expression of the indicated shRNAs targeting Apc. **g**, Competitive proliferation assays of MLL-AF9;Nras^{G12D} leukaemia cells transduced with pLMP constructs expressing shRNAs targeting Apc, in combination with JQ1 (50 nM) or DMSO over time (continuation from Fig. 4d). **h**, Top, bioluminescent imaging of mice transplanted with 1×10^5 MLL-AF9;Nras^{G12D} leukaemia cells expressing constitutively active Ctnnb1 (Ctnnb1.4x). Treatment with JQ1 (50 mg kg⁻¹ per day) or DMSO carrier started at day 1 after injection. Bottom, Kaplan-Meier survival curves of control and JQ1-treated mice demonstrate decreased survival rates in mice treated with JQ1 ($n = 5$). Statistical significance was calculated using the log-rank test. **i**, Competitive proliferation assays of BCR/ABL^{p210};p53^{-/-} and MLL/ENL;Nras^{G12D} leukaemia cells transduced with constitutively active Ctnnb1 (Ctnnb1.4x) or empty vector control in the absence or presence of JQ1. Measurements started 4 days after transduction together with JQ1 treatment (50 nM). Shown is the fraction of GFP⁺ cells (relative to initial) over time.



Extended Data Figure 9 | Wnt signatures are generally associated with BET resistance and Wnt activation drives resistance of pancreatic cancer models.

The BETi_Resistance_Rathert signature was defined by combining the core enriched genes from the two significant Wnt gene sets (KEGG_WNT_SIGNALING and ST_WNT_SIGNALING) in murine resistant AML, filtered for significant upregulation (DESeq padj < 0.1). These were combined with the Wnt-associated genes found differentially expressed in resistant human leukaemia cell lines and primary patient samples, resulting in a total of 26 genes. **a**, Left, microarray expression data of all 26 signature genes was curated from the Cancer Cell Line Encyclopedia (CCLE)²⁸ and normalized to the geometric mean of each individual gene throughout all samples (relative expression). The sum of the relative expression of all genes (resistance index) was plotted for all CCLE cell lines showing a $GI_{50} > 450$ nM (resistant, 54 total) or a $GI_{50} < 150$ nM (sensitive, 55 total) based on our sensitivity profiling (statistical significance was determined using Student's *t*-test). Right, gene set enrichment analysis plot comparing the expression of 26 signature genes associated with JQ1 resistance across 54 resistant ($GI_{50} > 450$ nM) and 55 sensitive cell lines ($GI_{50} < 150$ nM) available from CCLE. **b**, Left, as in **a** the resistance index was calculated as the sum of relative expression values of all 26 signature genes, which were based on RPKM extracted from an independent RNA-seq data set²⁷. Plotted are all 49 resistant ($GI_{50} > 450$ nM) and all 50 sensitive ($GI_{50} < 150$ nM) cell lines analysed in both sensitivity profiling and RNA-seq²⁷

(statistical significance determined using Student's *t*-test). Right, gene set enrichment analysis plot comparing the expression of 26 signature genes associated with JQ1 resistance across 49 resistant ($GI_{50} > 450$ nM) and 50 sensitive cell lines ($GI_{50} < 150$ nM) available from ref. 27. **c**, Competitive proliferation assays of murine pancreatic adenocarcinoma (KRPC2) cells transduced with pLEPC constructs expressing potent validated shRNAs targeting Apc, cultured in the presence of JQ1 (800 nM) or DMSO. Shown is the relative number of mCherry⁺ cells over time, relative to initial. **d**, Cell viability of murine KRPC2 was determined following 5 days of treatment with pyrivinium and/or JQ1 as indicated ($n = 3$, mean \pm s.e.m., statistical significance determined using Student's *t*-test). The combination index (CI) for drug combinations was calculated using the CompuSyn software and percentage inhibition (fraction affected, Fa) resulting from combined action of the two drugs versus effects of either drug alone. CI values < 1.0 indicate synergism of the two agents. **e**, As in **d**, cell viability was determined for human ASPC-1 pancreatic cancer cells following 5 days of treatment with pyrivinium and/or JQ1 as indicated ($n = 3$, mean \pm s.e.m., statistical significance determined using Student's *t*-test). The combination index for drug combinations was obtained using percentage inhibition (fraction affected, Fa) resulting from combined action of the two drugs versus effects of either drug alone.



Extended Data Figure 10 | Expression analysis of Wnt-associated genes in primary AML patient samples. **a**, Determination of JQ1 response profiles in 12 primary AML patient samples. Sensitivity was determined using 3H-thymidine uptake across different JQ1 concentrations. **b**, qRT-PCR analysis of mRNA levels of additional Wnt-associated genes in primary human leukaemia samples relative to *GAPDH* (continuation of Fig. 4g). Patient groups with low JQ1 IC₅₀ (<200 nM, blue dots) were compared to patients with high IC₅₀ (>500 nM, red dots). Statistical significance was determined using a

Student's *t*-test. **c**, Definition of a JQ1 resistance index. Expression of *HOXB4*, *TCF4* and *CCND2* in each primary AML patient sample was normalized to the geometric mean of all samples. The sum of these relative expression values of all three genes were added up to a resistance index, which was plotted in comparison to the JQ1 IC₅₀ of each sensitive (IC₅₀ < 200 nM, blue dots) and resistant (IC₅₀ > 500 nM, red dots) AML patient sample. Two-tailed Pearson correlation coefficient *r* and *P* value are shown.

Crystal structures of a double-barrelled fluoride ion channel

Randy B. Stockbridge¹, Ludmila Kolmakova-Partensky¹, Tania Shane¹, Akiko Koide², Shohei Koide², Christopher Miller^{1,3} & Simon Newstead⁴

To contend with hazards posed by environmental fluoride, microorganisms export this anion through F^- -specific ion channels of the Fluc family^{1–4}. Since the recent discovery of Fluc channels, numerous idiosyncratic features of these proteins have been unearthed, including strong selectivity for F^- over Cl^- and dual-topology dimeric assembly^{5,6}. To understand the chemical basis for F^- permeation and how the antiparallel subunits convene to form a F^- -selective pore, here we solve the crystal structures of two bacterial Fluc homologues in complex with three different monobody inhibitors, with and without F^- present, to a maximum resolution of 2.1 Å. The structures reveal a surprising ‘double-barrelled’ channel architecture in which two F^- ion pathways span the membrane, and the dual-topology arrangement includes a centrally coordinated cation, most likely Na^+ . F^- selectivity is proposed to arise from the very narrow pores and an unusual anion coordination that exploits the quadrupolar edges of conserved phenylalanine rings.

The fluoride anion, ubiquitous in the aqueous biosphere throughout evolutionary time, is a xenobiotic inhibitor of essential phosphoryl-transfer enzymes⁷. Unicellular organisms directly exposed to environmental F^- counteract the toxicity of the anion through the action of F^- -exporting membrane transport proteins that keep cytoplasmic F^- below inhibitory levels^{1–4,8}. Two recently discovered phylogenetically unrelated families of F^- exporters carry out this task: the $CLC^F F^-/H^+$ antiporters, a strictly bacterial clade of the CLC superfamily of anion transporters, and small-membrane proteins of the Fluc family (also known as CrcB or FEX). Fluc genes are also found in plants, fungi, primitive marine chordates, and sponges, but not in mammals. We recently established^{5,6} that Fluc proteins are ion channels with two unusual properties: an exceedingly high specificity ($>10^4$) for F^- over Cl^- , and a dual-topology dimeric architecture, in which the two subunits forming the active channel associate in antiparallel transmembrane ori-

entation. Dual-topology dimeric construction is known in small multidrug transporters^{9–11}, but has not been previously observed in ion channels. These inferences regarding function and structure of Flucs provoke fundamental questions about their mechanisms, such as (1) how the channel achieves such extreme selectivity for F^- , arguably the highest selectivity of any ion channel to our knowledge; (2) whether the protein contains a single pore on the subunit interface or two pores, one in each subunit, as in $CLC Cl^-$ channels¹²; and (3) whether the channel homodimer is symmetrical, or the two subunits adopt different conformations, as in the multidrug resistance transporter EmrE^{9,11}.

In previous work, the antiparallel transmembrane topology of Fluc channels was intimated by the distribution of positively charged residues in Fluc sequences¹³, was strongly suggested by crosslinking and functional reconstitution⁵, and was established definitively by two-sided block of single channels by ‘monobodies’, engineered proteins selected as high-affinity binding partners from combinatorial libraries⁶. We used these monobodies in crystallization trials to form complexes with a Fluc channel from *Bordetella pertussis*, denoted Bpe⁵. Crystals could not be grown unless monobodies were present; however, diffraction to 3.6 Å Bragg spacing was obtained with the monobody Mb(Bpe-S7) (ref. 6), hereafter denoted S7. A structure was solved, with phases initially determined by single-wavelength anomalous diffraction (SAD) of Bpe labelled with Hg at a unique cysteine residue, and improved with Hg-labelled selenomethionine-substituted samples (Extended Data Table 1). A view of the crystal lattice highlights the importance of the monobodies as crystallization chaperones, which exclusively mediate crystal contacts in all structures presented here (Extended Data Fig. 1).

Although devoid of F^- ions and at low resolution, the Bpe-S7 structure reveals the overall architecture of the channel (Fig. 1a). Bpe is an antiparallel homodimer in which each subunit consists of

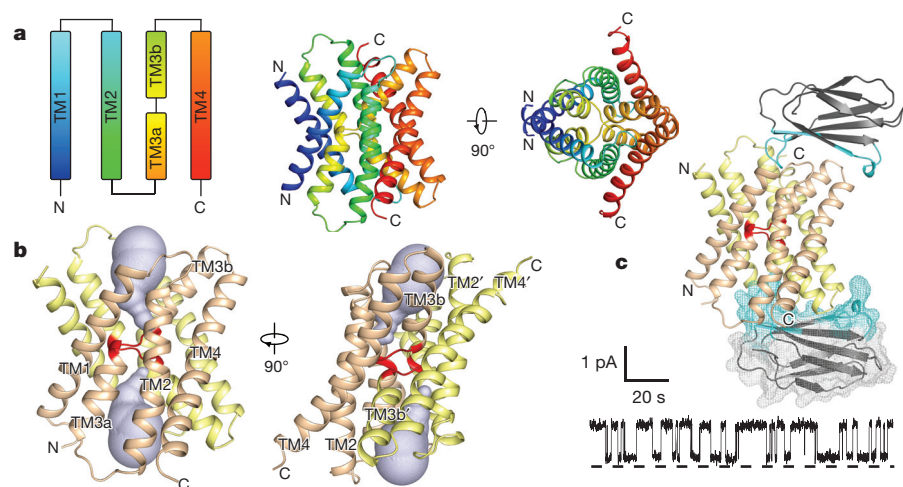


Figure 1 | Bpe-S7 structure. **a**, Bpe homodimer, viewed parallel (centre structure) and perpendicular (right structure) to the membrane. Colouring is as indicated in the transmembrane helix schematic on the left. **b**, Through-membrane views, at the same orientation as in **a**, indicating the TM3 break (red) and the aqueous volumes within the vestibules (blue). **c**, Bpe with S7 monobodies bound. Variable sequences of the monobodies (cyan) are shown with ribbon or mesh representation. A single-channel Bpe recording showing S7 (500 nM) binding and dissociation is shown below. Zero-current level is indicated by the dashed line.

¹Department of Biochemistry, Howard Hughes Medical Institute, Brandeis University, Waltham, Massachusetts 02454, USA. ²Department of Biochemistry and Molecular Biology, University of Chicago, Chicago, Illinois 60637, USA. ³Department of Physiology, Anatomy, and Genetics, University of Oxford, Oxford OX1 3QU, UK. ⁴Department of Biochemistry, University of Oxford, Oxford OX1 3QU, UK.

four transmembrane helices (denoted TM), with an overall fold that is novel among membrane proteins. The 1,700 Å² dimer interface is almost completely membrane-embedded. The third, highly conserved helix (TGXXXGLTTFSTFXXE, in which X denotes any amino acid) is broken into two halves, TM3a and TM3b, by a six-residue non-helical segment located roughly at the centre of the membrane. These two segments, one from each subunit, cross each other near the two-fold axis of the channel running parallel to the membrane plane. The channel is hourglass-shaped, with wide vestibules symmetrically opening to two aqueous solutions (Fig. 1b) separated by a solid plug of protein 10–15 Å thick. A conspicuous universally conserved TM1 arginine residue (R23) protrudes into each vestibule, suggestive of an electrostatic lure for F[−]. No aqueous pore connecting the vestibules is visible in this low-resolution structure.

The channel is capped on both ends by the S7 monobody (Fig. 1c). This monobody was selected from a library designed to target convex protein surfaces¹⁴, and indeed its interaction surface, consisting largely of the residues diversified in the library, wraps around a protrusion formed from the TM1–TM2 and TM3b–TM4 connecting loops of the channel. An eight-residue loop on the monobody plunges deeply into each vestibule, contacting the channel mainly via side chains. Channel–monobody interactions are mostly hydrophobic, aromatic, and hydrogen bonded, the paucity of salt-bridges rationalizing the rather weak ionic strength dependence of monobody binding¹⁵. Most of the aqueous-exposed surface of the channel is covered by monobody, consistent with S7 block of Bpe seen in single-channel recordings (Fig. 1c).

It is tempting to imagine that a central pore connects the two vestibules in an unseen ‘open’ conformation. But the low resolution of the structure and the absence of F[−] preclude identification of the ion-permeation pathway. We therefore attempted *in meso* crystallization in the hope of identifying bound F[−] ions. Crystals diffracting to 2.1 Å in the presence of 20 mM F[−] were obtained with a different monobody, L2, which is also a blocker (Extended Data Fig. 2). Structures solved by molecular replacement (Extended Data Table 1) again show the channel with a monobody on each end (Extended Data Figs 1 and 2). The backbone conformation of the channel is identical to that in the lower-resolution structure (C α root mean squared deviation (r.m.s.d.), 0.4 Å), and L2, although binding in a different orientation than S7, also extends a long loop of 8–10 Å into the vestibule, occluding much of the channel’s water-exposed surface.

This higher-resolution structure reveals five intriguing electron densities (Fig. 2a and Extended Data Fig. 3). First, a prominent density resides in the centre of the plug separating the vestibules, precisely on the homodimer’s two-fold non-crystallographic axis (Fig. 2b). We

identify this as a Na⁺ ion on the basis of its coordination by four backbone carbonyl groups from residues in each subunit associated with the conserved TM3 break (G77 and T80). This coordination is inconsistent with a F[−] ion, a water, a divalent metal, or a K⁺ ion^{16,17}. Although coordination by only four oxygen ligands is uncommon for Na⁺, it is nevertheless seen in ~15% of Na⁺-binding sites in the protein database¹⁸. This deeply buried cation could not exchange with aqueous solution if the plug remained intact during functional activity, and indeed, Bpe channels with familiar behaviour are readily recorded in solutions with Na⁺ completely substituted by *N*-methyl glucamine (Extended Data Fig. 4). We propose that the ion is an important structural component incorporated irreversibly upon dimer assembly.

A second notable detail in the Bpe–L2 structure is a set of four electron densities located in crevices between TM2, TM3b, and TM4 near the periphery of the channel, distant from the vestibules and the central plug (Fig. 2a, c). We provisionally identify these as F[−] ions, labelled F1 and F2 in non-crystallographic-symmetry-related pairs, according to their distinct chemical environments. The putative liganding atoms embracing these densities are consistent with a halide coordination shell. In particular, the surround is composed of electropositive side chains, which would engage the strong H-bond-accepting tendency of the F[−] ion. Prominent among these are a strongly conserved asparagine (N43) in TM2, and two conserved serines (S108 and S112) in TM4. In addition, two pairs of conserved phenylalanine rings (F82 and F85) near the TM3 break approach these densities in a side-on orientation that presents the electropositive carbons of the quadrupolar ring to the F[−] ion (Fig. 2d). These four aromatic rings appear to be mutually stabilized in a notable ‘box’ assembly. Edge-on aromatic liganding of anions is rare but not unprecedented in proteins¹⁹, and F[−] coordination by aromatic edges appears in many small-molecule structures²⁰. This type of coordination is reminiscent of the phenylalanine rings of a proposed Cl[−]-binding site in the bestrophin channel²¹. With a deficit of H-bond acceptors, the coordination shells observed here would be chemically inimical to ordered waters, which cannot be distinguished from F[−] based on X-ray scattering alone.

The two F[−] ions of each pair lie in a vertical line tilted slightly off normal to the membrane plane, possibly marking a narrow permeation pathway. If these densities do indeed represent F[−], then their positions lead to a surprising conclusion: the Bpe channel contains two pores running in antiparallel orientation along opposite sides of the dimer, rather than a single central pore connecting the vestibules through the plug. Two-pore behaviour is not apparent in single-channel recordings as both are nearly always open⁵. The structures here would represent the monobody-blocked state similar to that observed electrophysiologically^{6,15}.

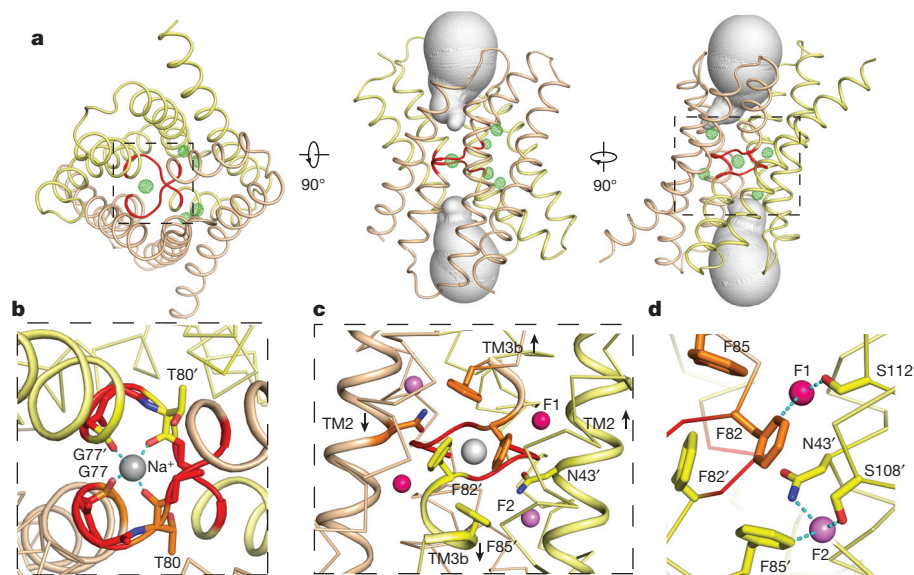


Figure 2 | Bpe–L2 structure. **a**, Bpe homodimer with mF_o – DF_c electron density map contoured at 4σ (green), viewed from solution (left) and the membrane (centre and right). Aqueous volumes of the vestibules are shown (grey). Dashed boxes in right- and left-most structures indicate zoomed regions in **b** and **c**, respectively. **b**, Na⁺ coordination sphere, indicated by blue dashed lines, viewed from solution. TM3 is represented as a cartoon. **c**, The same view at 90°, showing Na⁺, F[−] ions, N43, and the phenylalanine box (F82 and F85). Cartoons indicate TM2 and TM3. **d**, F[−] coordination shells, indicated by blue dashed lines, with F82, S112 (at F1), N43, S108, and F85 (at F2) shown as sticks. Stereo images are shown in Extended Data Fig. 3.

Our reading of the structure as a double-barrelled channel depends crucially on identifying these four densities as F^- ions; however, we do not consider the above evidence sufficient to accept such an unusual picture as firmly established. Accordingly, two additional experimental approaches were pursued. First, a structure was determined for a different Fluc homologue, Ec2, complexed with a monobody S9 (Extended Data Fig. 5). This homologue of only 33% identity shows similar electrophysiological behaviour to Bpe, and is blocked by S9 at nanomolar concentrations⁶. Ec2–S9 crystals diffracting to 2.6 Å were grown from detergent in the presence of F^- , and to avoid model bias the structure was solved using SAD phasing with selenomethionine-labelled protein (Extended Data Table 1). The Ec2 and Bpe folds are identical ($C\alpha$ r.m.s.d., 0.6 Å), with the inferred Na^+ density appearing in equivalent locations. Two strong difference densities lie at precisely the same locations as the F1 ions in Bpe, coordinated identically (Fig. 3a). The appearance of these densities, in a separate homologue under very different crystallization conditions, strengthens our hypothesis that the four densities in Bpe represent F^- ions. Additional densities also appear in Ec2 in the general vicinity of the F2 ions in Bpe, but at this lower resolution and without supporting experimental evidence, these densities in Ec2 cannot be unambiguously assigned.

The chemical nature of the crevices housing the densities makes sense for narrow diffusion pathways that are welcoming to F^- ions in both homologues (Fig. 3b). In particular, the crevice-facing surface of TM4 is lined with H-donating side chains, as manifested by every fourth residue in the sequences of Bpe (Y104, S108, S112, and T116) or Ec2 (S102, H106, S110, and T114). These particular residues are only modestly conserved among Fluc channels, but H-bond donors consistently appear here throughout the family. These could plausibly contribute to a polar track, along which largely dehydrated F^- ions move across the membrane. Because these pathways are extremely narrow, protein dynamics may be necessary to allow F^- permeation, and the monobodies might force a conformation in which the two pores are less ‘open’ than in the fully conducting state.

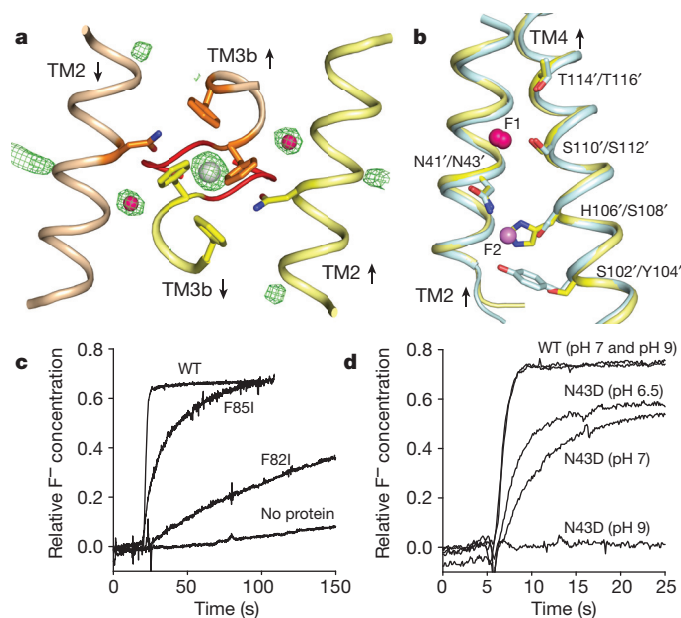


Figure 3 | Identification of F^- ions. **a**, Ec2–S9 F^- -binding region, analogous to Fig. 2c, with $mF_o - DF_c$ map contoured at 4.8σ and ion positions shown. The phenylalanine box and N41 are shown as sticks. Weak densities in F2 regions could not be unambiguously assigned. **b**, Overlay of TM2 and TM4 from Bpe–L2 (cyan) and Ec2–S9 (yellow) structures, showing F^- ions and polar track. **c**, **d**, Functional consequences of mutating F^- -coordinating residues. F^- efflux from Bpe-reconstituted proteoliposomes, initiated with valinomycin and monitored with a F^- electrode, reported relative to total F^- trapped in liposomes. WT, wild type.

We next examined the functional consequences of mutating each of the three conserved Bpe residues coordinating the F^- densities: F82, F85, and N43. We did not observe single-channel activity in electrical recordings, but more sensitive ‘anion-dump’ experiments^{22,23} reveal notable changes in F^- permeation (Fig. 3c). In these experiments, Bpe-reconstituted liposomes loaded with KF are suspended in low- F^- solutions, and the rate of passive F^- efflux is followed electrochemically. To eliminate the aromatic quadrupole, the conserved phenylalanine residues were mutated individually to isoleucine. For F85I and F82I, efflux of F^- is two and three orders of magnitude slower, respectively, than for wild type ($\sim 3 \times 10^5 s^{-1}$; ref. 15) (Extended Data Table 2). These mutations preserve the integrity of the channel, as F^- efflux is 50–80% blocked by 6 μM of monobody (Extended Data Fig. 6). Strong selectivity against Cl^- is observed in parallel experiments with this halide.

The conserved asparagine was substituted to alter or remove H-bonding capability (N43S and N43A), or to place an isosteric carboxylate at this position (N43D). The first two mutants were biochemically intractable, but N43D produced stable protein. Under standard conditions at pH 7, N43D supports robust F^- -selective efflux (Fig. 3d and Extended Data Table 2). We had envisioned that an anionic carboxylate at this position would prevent F^- entry into the channel. It is possible, however, that the pK_a of this group is perturbed upwards by its local environment, so that at neutral pH conditions the carboxyl acts as a protonated surrogate for the N43 amide. The N43D mutant was therefore examined at several pH values. In stark contrast to the pH-independent activity of wild-type Bpe and the F85I mutant (Extended Data Fig. 5), F^- efflux in N43D falls with increasing pH and is extinguished at pH 9 (Fig. 3d), verifying a key role of the H-bond-donating N43 side chain in F^- conduction.

These mutagenic manipulations of F^- permeation add mechanistic evidence for assigning F^- ions to the densities in question. This inference points to the conclusion that Flucs are double-barrelled F^- channels, with the observed F^- ions marking the ion-selective pathways. The two pathways are not segregated to each subunit as in CLC channels²⁴; instead, each pore comprises side chains from TM2, TM3b, and TM4 of one subunit plus the TM3-break phenylalanine from the opposing subunit. Although unexpected, this idea does not clash with any electrophysiological experiments, and double-pore assembly was cited previously as a possible, albeit unlikely, architecture consistent with the functional behaviour of the channel⁵.

Two-pore assembly neatly accounts for evolutionary drift in eukaryotic Fluc channels, all of which consist of an inverted repeat of two homologous Fluc domains fused into a single polypeptide^{4,5}. Alignments (Extended Data Fig. 7) show that ‘pore 2’, where most residues arise from the carboxy-terminal domain, retains the strict sequence conservation typical of the homodimeric bacterial Flucs, whereas the equivalent residues of ‘pore 1’, mostly in the amino-terminal domain, are far less conserved, notably along the TM4 polar track. This pattern is further confirmed for the two conserved phenylalanines that, from the same domain, contribute to alternate pores. Thus, in pore 2, the strongly conserved equivalents of F82 project from the N-terminal domain, and those of F85 from the C-terminal domain, in contrast to their poorly conserved residue-counterparts contributing to pore 1. These sequence-based considerations suggest that pore 2 alone fulfils the F^- -export function in eukaryotic Flucs. Recent experiments reinforce this idea by showing that mutations of several C-terminal domain residues in pore 2 produce F^- hypersensitivity in yeast, whereas mutations of the equivalent residues of N-terminal domain, pore 1 are relatively harmless⁴. These features chronicle an evolutionary lineage of gene duplication, fusion, and finally functional degradation of a redundant pore by genetic drift in the eukaryotic homologues.

Other inferences emerging from these structures will require further testing to confirm or refute. First, a buried Na^+ ion occupying a unique position on the two-fold axis invites us to view this cation as an intrinsic structural element stabilizing the dimer interface. Second, all

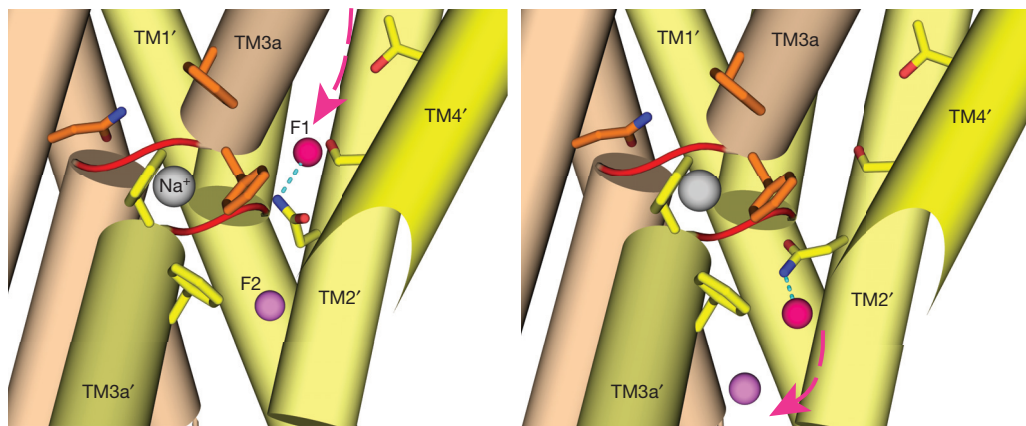


Figure 4 | Model of F^- movement by Fluc 'channsporter' mechanism. Each panel shows the proposed successive steps of F^- ion electrodiffusion through a two-ion pore, facilitated by the concerted rotation of the N43 side chain. The lower half of the TM4 polar track is removed for clarity.

four F^- ions observed in Bpe probably occupy the channel simultaneously, given their high occupancies (>80%) in the 2.1 Å structure, with B-factors matched to those in their coordination shells. The channel might therefore display multi-ion conduction phenomena akin to those long-known in K^+ channels^{25,26}. Third, the strong F^- selectivity of the channel may arise from the narrow bore of the permeation pathway, which would exclude Cl^- ions while permitting the smaller F^- ions to enter. However, it is unclear why F^- would enter this confined space, and how the protein compensates for the high energy of dehydrating F^- . We note that many of the coordinating groups are H-bond donors, able to satisfy the H-bonding proclivity of the F^- ion. While the unusual edge-on coordination by conserved phenylalanine rings is chemically intriguing, the energetic contributions of these interactions have not been established; nevertheless, in light of the conservation of these residues, we speculate that this short-range quadrupolar interaction contributes to F^- recognition and permeation in an essential way. The pore, although narrow, is lined with H-bonding residues and so could provide a polar conduit for transport of the ion across the membrane span.

A final point concerns the mechanism of electrodiffusive F^- transport through these oddly fashioned pores. The crucial role of N43 in permeation in Bpe, and the confined crevice in which it resides, lead us to conjecture that F^- moves along the pore concomitant with a rotameric switch of this side chain, such that the amide nitrogen remains H-bonded as the anion moves along the pore (Fig. 4). Thus, the conduction mechanism we propose here would be subtly distinct from classic diffusion through a fixed, water-filled channel. Instead, it would incorporate a central feature of membrane transporters: substrate transport coupled to concerted movement of the protein. An asparagine side-chain rotation could easily occur on the conduction timescale of microseconds, but formally this picture is a nuanced mix of electrodiffusion and configurational change, and so can be termed a 'channsporter' mechanism.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 15 April; accepted 20 July 2015.

Published online 7 September 2015.

- Baker, J. L. *et al.* Widespread genetic switches and toxicity resistance proteins for fluoride. *Science* **335**, 233–235 (2012).
- Li, S. *et al.* Eukaryotic resistance to fluoride toxicity mediated by a widespread family of fluoride export proteins. *Proc. Natl Acad. Sci. USA* **110**, 19018–19023 (2013).
- Ji, C., Stockbridge, R. B. & Miller, C. Bacterial fluoride resistance, Fluc channels, and the weak acid accumulation effect. *J. Gen. Physiol.* **144**, 257–261 (2014).
- Smith, K. D. *et al.* Yeast FEX1 is a constitutively expressed fluoride channel in with functional asymmetry of its two homologous domains. *J. Biol. Chem.* **290**, 19874–19887 (2015).
- Stockbridge, R. B., Robertson, J. L., Kolmakova-Partensky, L. & Miller, C. A family of fluoride-specific ion channels with dual-topology architecture. *eLife* **2**, e01084 (2013).
- Stockbridge, R. B., Koide, A., Miller, C. & Koide, S. Proof of dual-topology architecture of Fluc F^- channels with monobody blockers. *Nature Commun.* **5**, 5120 (2014).

- Adamek, E., Pawlowska-Goral, K. & Bober, K. *In vitro* and *in vivo* effects of fluoride ions on enzyme activity. *Ann. Acad. Med. Stetin.* **51**, 69–85 (2005).
- Stockbridge, R. B. *et al.* Fluoride resistance and transport by riboswitch-controlled CLC antiporters. *Proc. Natl Acad. Sci. USA* **109**, 15289–15294 (2012).
- Ubarretxena-Belandia, I., Baldwin, J. M., Schuldiner, S. & Tate, C. G. Three-dimensional structure of the bacterial multidrug transporter EmrE shows it is an asymmetric homodimer. *EMBO J.* **22**, 6175–6181 (2003).
- Schuldiner, S. EmrE, a model for studying evolution and mechanism of ion-coupled transporters. *Biochim. Biophys. Acta* **1794**, 748–762 (2009).
- Morrison, E. A. *et al.* Antiparallel EmrE exports drugs by exchanging between asymmetric structures. *Nature* **481**, 45–50 (2011).
- Miller, C. Open-state substructure of single chloride channels from *Torpedo* electroplax. *Philos. Trans. R. Soc. Lond. B* **299**, 401–411 (1982).
- Rapp, M., Granseth, E., Seppala, S. & von Heijne, G. Identification and evolution of dual-topology membrane proteins. *Nature Struct. Mol. Biol.* **13**, 112–116 (2006).
- Koide, A., Wojcik, J., Gilbreth, R. N., Hoey, R. J. & Koide, S. Teaching an old scaffold new tricks: monobodies constructed using alternative surfaces of the FN3 scaffold. *J. Mol. Biol.* **415**, 393–405 (2012).
- Turman, D. L., Nathanson, J. T., Stockbridge, R. B., Street, T. O. & Miller, C. Two-sided block of a dual-topology F^- channel. *Proc. Natl Acad. Sci. USA* **112**, 5697–5701 (2015).
- Harding, M. M. Metal-ligand geometry relevant to proteins and in proteins: sodium and potassium. *Acta Crystallogr. B* **58**, 872–874 (2002).
- Varma, S., Sabo, D. & Rempe, S. B. K^+/Na^+ selectivity in K channels and valinomycin: over-coordination versus cavity-size constraints. *J. Mol. Biol.* **376**, 13–22 (2008).
- Lev, B. R. B. & Noskov, S. Y. in *Encyclopedia of Metalloproteins* (eds Kretsinger, R. H. *et al.*) 2112–2118 (Springer, 2013).
- Philip, V. *et al.* A survey of aspartate-phenylalanine and glutamate-phenylalanine interactions in the protein data bank: searching for anion- π pairs. *Biochemistry* **50**, 2939–2950 (2011).
- Allen, F. H. The Cambridge Structural Database: a quarter of a million crystal structures and rising. *Acta Crystallogr. B* **58**, 380–388 (2002).
- Kane Dickson, V., Pedi, L. & Long, S. B. Structure and insights into the function of a Ca^{2+} -activated Cl^- channel. *Nature* **516**, 213–218 (2014).
- Walden, M. *et al.* Uncoupling and turnover in a Cl^-/H^+ exchange transporter. *J. Gen. Physiol.* **129**, 317–329 (2007).
- Brammer, A. E., Stockbridge, R. B. & Miller, C. F. F^-/Cl^- selectivity in CLCF-type F^-/H^+ antiporters. *J. Gen. Physiol.* **144**, 129–136 (2014).
- Weinreich, F. & Jentsch, T. J. Pores formed by single subunits in mixed dimers of different CLC chloride channels. *J. Biol. Chem.* **276**, 2347–2353 (2001).
- Hille, B. & Schwarz, W. Potassium channels as multi-ion single-file pores. *J. Gen. Physiol.* **72**, 409–442 (1978).
- Morais-Cabral, J. H., Zhou, Y. & MacKinnon, R. Energetic optimization of ion conduction rate by the K^+ selectivity filter. *Nature* **414**, 37–42 (2001).

Acknowledgements We are grateful to J. Parker for critically reading the manuscript; to D. Turman, A. Lajoie, M. Pham, and K. Piasta for pilot experiments; to S. Strobel for sharing results of unpublished experiments; and to B. Foxman for help with the Cambridge Structural Database. This work was supported in part by a Wellcome Trust Investigator Award 102890/Z/13/Z and National Institutes of Health (NIH) grants R01-GM107023 and U54-GM087519. R.B.S. was supported by NIH grant K99-GM-111767. C.M. is grateful to F. Ashcroft and Lincoln College, Oxford, for hosting a Newton-Abraham Visiting Professorship.

Author Contributions R.B.S., S.N. and C.M. performed experiments, analysed the data and wrote the paper. L.K.-P. and T.S. performed experiments. S.K. and A.K. provided monobody clones and wrote the paper.

Author Information Atomic coordinate files have been uploaded to the Protein Data Bank (PDB) with accession codes 5A40, 5A41 and 5A43 for the Bpe-S7, Bpe-L2 and Ec2-S9 complexes, respectively. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to S.N. (simon.newstead@bioch.ox.ac.uk).

METHODS

No statistical methods were used to predetermine sample size.

Preparation of crystals. Expression, purification, and reconstitution of Fluc channels were performed as previously described in detail^{5,6}. In the final purification step, Fluc protein was collected from a S200 size-exclusion column equilibrated in 100 mM NaF (or NaCl for zero- F^- preparations), 10 mM HEPES, pH 7.0, and 5 mM *n*-decyl- β -D-maltoside (DDM). Bpe constructs carried two functionally neutral mutations to enhance expression (R29K and E94S) or for Hg labelling (R29K and E94C). Hg labelling was achieved by incubating Bpe with a threefold molar excess of Hg(II) acetate for 30 min between the co-affinity and size-exclusion columns. Ec2 constructs bore a single functionally neutral expression-enhancing mutation (R25K), and, for selenomethionine incorporation, an additional methionine was introduced (A51M) to enhance phasing power. The C-terminal His₆ tag was removed from Bpe by treatment with lysine endoproteinase C (Roche)⁵, but was left on Ec2. Fluc protein was typically reconstituted into liposomes at low density (0.1–0.2 μ g of protein per mg of lipid). For single-channel recording, liposomes were fused into planar lipid bilayers in symmetrical solutions of 300 mM NaF, 15 mM MOPS, pH 7.0, and channels were recorded at 200 mV holding voltage^{5,15}. Monobodies were expressed in *Escherichia coli* and purified as previously described⁶. N-terminal His₆ tags were removed while bound to talon beads by 16-h treatment with tobacco etch virus (TEV) protease also carrying a His tag; monobodies with cleaved His tags were eluted from the affinity column with 150 mM NaCl and 40 mM Tris-HCl, pH 7.5. For the final purification step, the preparation was passed over a S75 size-exclusion column in 100 mM NaF (or NaCl) and 10 mM HEPES, pH 7. Monobodies were used immediately for crystallization or stored in frozen aliquots for channel-blocking experiments. For crystallization from detergent micelles, Fluc protein in solution containing 5 mM DDM was concentrated to 10 mg ml⁻¹, a step that concentrates the detergent 5–10-fold. Monobody solution (10 mg ml⁻¹) was supplemented with 4 mM DDM immediately before mixing with channel in a 1.2:1 molar ratio. This protein solution was then mixed with an equal volume of crystallization solutions (0.5 μ l for sitting drops in 96-well plates or 1 μ l for hanging drops in 24-well plates). Bpe-S7 crystals grew in 3–5 days in crystallization solutions of 36–41% (w/v) polyethylene glycol monomethyl ether 550, 0.2 M MgCl₂ or CaCl₂, and 0.1 M Tris, pH 8.5–8.9. Ec2-S9 crystals grew in 10–14 days in crystallization solutions containing 28–32% (w/v) polyethylene glycol monomethyl ether 550, 0.05 M LiNO₃, and 0.1 M *N*-(2-acetamido)iminodiacetic acid, pH 6.0–6.7. Crystals were frozen in liquid nitrogen for data collection. For lipidic cubic phase crystallization, Fluc protein concentrated to 10 mg ml⁻¹ as above was dialysed overnight to reduce the DDM concentration to 10 mM. This was then mixed with monobody solution (10 mg ml⁻¹, with 4 mM DDM) in a 1:1.2 molar ratio. The protein-laden mesophase was prepared by homogenizing 9.9 monoacylglycerol (monoolein) lipid with protein solution (10 mg ml⁻¹) at a weight ratio of 1:1.5 (protein:lipid) using a coupled syringe mixing device at 20 °C (ref. 27). Crystallization trials were carried out at 19 °C in 96-well glass sandwich plates with 50 nl mesophase and 0.8 μ l precipitant solution using an *in meso* robot. Crystallization solutions consisted of 22–26% (v/v) polyethylene glycol dimethyl ether 550, and 0.1 M Na-citrate, pH 5.5, with or without 10 mM NaF. Surfboard-shaped crystals grew to a maximum size of 100 \times 50 \times 5 μ m in 5–10 days. Wells were opened using a tungsten-carbide glasscutter, and the crystals were collected using 50–100 μ m micromounts (MiTeGen). Crystals were snap-cooled directly in liquid nitrogen before data collection on the Diamond Light Source beamlines I24 or I04.

Anion efflux from liposomes. Efflux of F^- or Cl^- out of liposomes was followed electrochemically as described^{22,23}. Liposomes (10 mg ml⁻¹ lipid, 0.2–1 μ g protein per mg of lipid) loaded with 300 mM KF or KCl solutions were freeze-thawed for 3 cycles and then extruded 21 times through a 400-nm filter. Immediately before the assay, a 100 μ l sample was centrifuged through a 1.5-ml Sephadex column equilibrated with flux buffer (300 mM K-isethionate, 1 mM KF or KCl, and 25 mM HEPES, pH 7) and was diluted 20-fold into a stirred chamber containing 3.8 ml flux buffer. Halide concentration in the suspension was continuously monitored with a F^- or Cl^- electrode amplified through a pH meter and digitized at 5 Hz sampling frequency. Efflux was initiated by adding 1 μ M valinomycin, and after several minutes 30 mM octylglucoside was added to obtain the 100% efflux level. Efflux rates were calculated after calibration with 25 μ M additions of NaF or NaCl. For experiments with the Bpe N43D mutant, the flux buffer contained an additional 100 mM Na-isethionate and 25 mM CHES (*N*-cyclohexyl-2-aminoethanesulfonic acid) buffer. Single-channel block by monobodies was recorded in planar phospholipid bilayers exactly as described⁶.

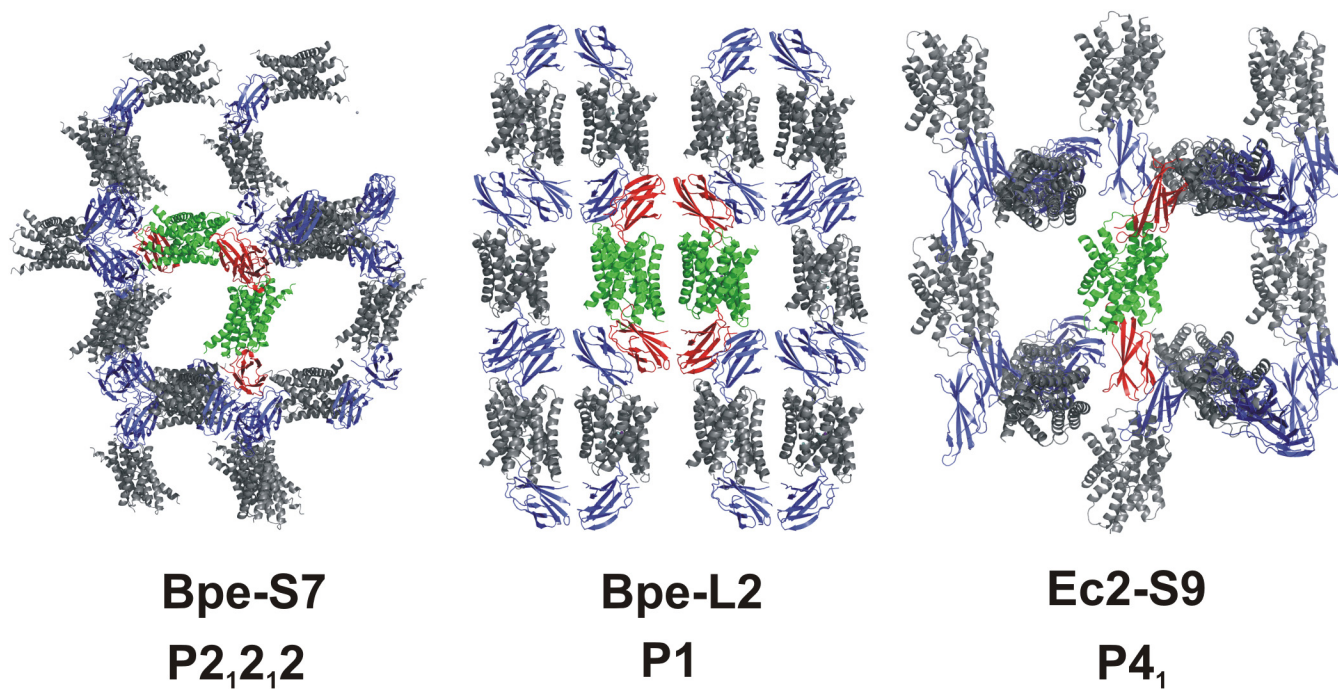
Structure determination. Diffraction data for Bpe-S7 were processed by the Xia2 pipeline²⁸ to XDS²⁹ and scaled using aimless³⁰. The space group was determined to be *P*₂,₁,₂ with two Bpe dimers and four S7 monobodies in the asymmetric

unit (Extended Data Fig. 1). A phasing strategy was devised that used pre-derivatization of Bpe mutated with a single cysteine residue (E94C) with Hg(II) acetate before crystallization (see above). None of the native crystals were isomorphous with the Hg-derivatized crystals (*R*_{iso} > 40%), despite having similar cell dimensions. Indeed, Hg-derivatized crystals were observed to diffract X-rays to slightly higher resolution than native crystals, therefore effort was directed at these samples for phasing and refinement. The four Hg sites were located using the SAD method as implemented in SHELX³¹ with the positions further refined and initial phases calculated using SHARP³² with solvent flattening in SOLOMON³³. To improve the phases, a second and third data set were also collected at the Se edge using both Hg- and seleno-L-methionine-derivatized protein and another Hg-derivatized data set, respectively (Extended Data Table 1). All 16 Se plus 4 Hg sites were located using SHELX and this data set was combined with the initial 3.6 Å Hg-derivatized data. Phases were substantially improved using SIRAS combining the three data sets in SHARP. We did not observe higher resolution diffraction in the native crystals, which typically gave diffraction between 3.6–3.8 Å. Our highest resolution data set with optimal scaling statistics was one of the Hg-derivatized crystals; we therefore used this data set for subsequent refinement of the model built into the experimental electron density maps calculated from SHARP (see below). For the Bpe-L2 crystals, data were similarly processed and scaled in space group *P*1. Phases were calculated using molecular replacement as implemented in Phaser³⁴, using the experimentally determined Bpe model and a homology model of the L2 monobody based on a previously determined structure of a loop-library monobody (PDB code: 3RZW). The unambiguous solution showed two Bpe homodimers and four L2 monobodies. The electron density maps clearly showed major differences in the selected variable regions of the monobody. For the Ec2-S9 crystals, the data were processed as above with space group *P*4₁. Phases were calculated using Se-SAD with eight Se sites, and processed as above. The experimental electron density maps were of high quality following phase extension to the highest resolution shell of 2.58 Å. Data were collected at Advanced Light Source beamlines 8.2.1 and 8.2.2, and Diamond Light Source beamlines I24 and I04.

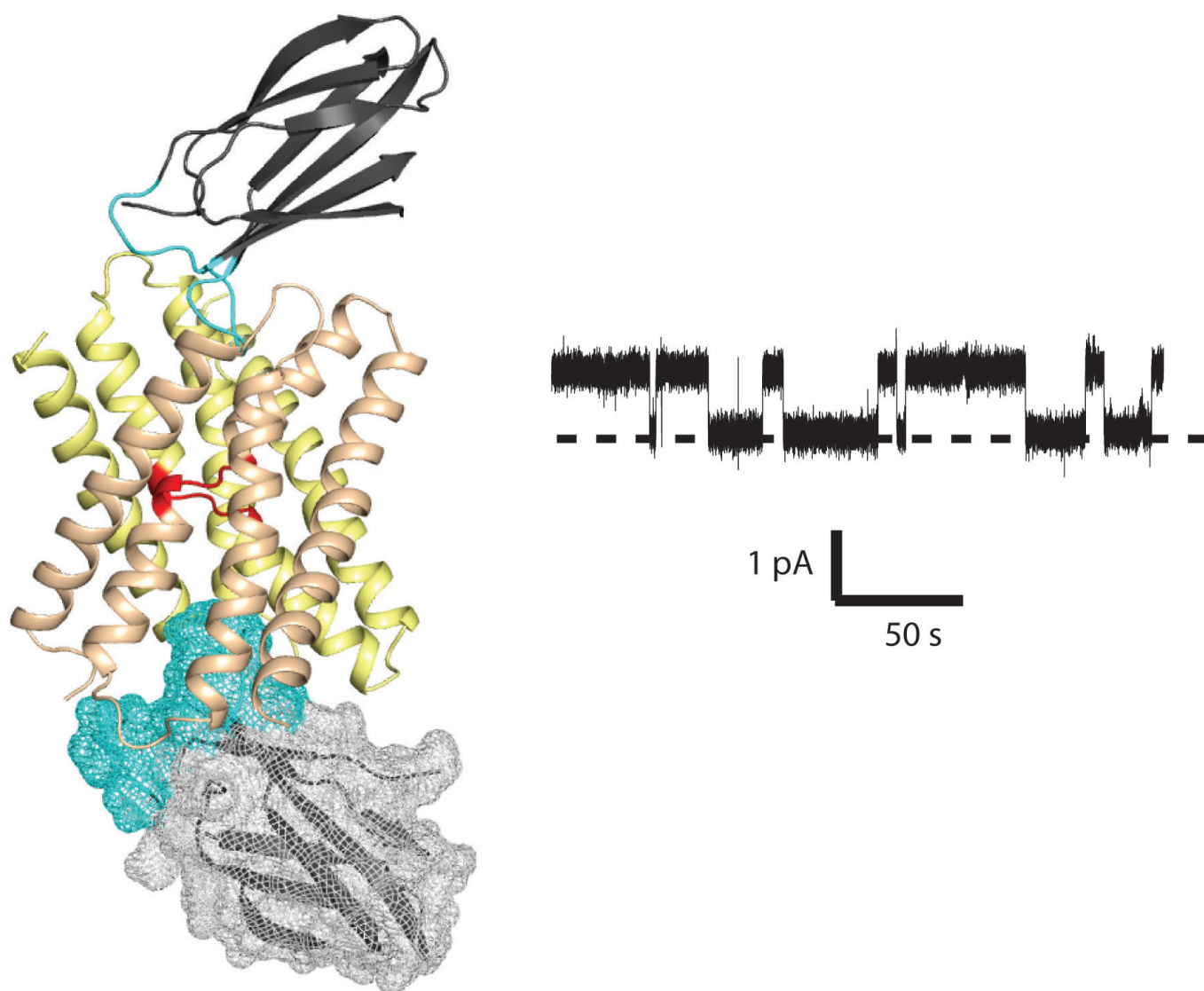
Model building and refinement. For the Bpe-S7 complex structure, a model for the channel was built into the experimental electron density maps calculated from SHARP using O³⁵ with σ A-weighted $2F_o - F_c$ and $mF_o - DF_c$ electron density maps. The S7 monobodies were initially built using a homology model based on a previously determined structure of a side-library monobody (PDB code: 4JEG). These models were placed into the experimental electron density maps using MolRep³⁶. The partial models were further cycled back into phase calculation in SHARP to improve the initial solvent envelope used for the solvent flipping procedure. The amino acid side chains were then built using the Se and Hg sites to determine the correct register. Refinement of the Bpe-S7 model was carried out in Refmac5 (refs 37, 38) against the highest resolution data set for these crystals, 3.6 Å, which came from one of the Hg-derivatized crystals used for phasing (Extended Data Table 1). No previous phase information was used during the refinement; however, refinement was improved following anisotropic truncation of the structure factors. To avoid biasing the model, non-crystallographic symmetry was not used except at the final round of refinement to improve model geometry. Model validation was carried out using the Molprobity server³⁹. The Ec2-S9 model was built directly into the experimental maps, using Se sites to ensure the correct register, and then monobodies were placed by molecular replacement using Phaser with a homology model based on S7. The Bpe-L2 model was built into the electron density maps calculated from Phaser following iterative rounds of structure refinement in PHENIX⁴⁰ and Refmac5. The structural model was revised in real space with Coot⁴¹.

27. Caffrey, M. & Cherezov, V. Crystallizing membrane proteins using lipidic mesophases. *Nature Protocols* **4**, 706–731 (2009).
28. Winter, G., Lobley, C. M. & Prince, S. M. Decision making in xia2. *Acta Crystallogr. D* **69**, 1260–1273 (2013).
29. Kabsch, W. Xds. *Acta Crystallogr. D* **66**, 125–132 (2010).
30. Evans, P. R. An introduction to data reduction: space-group determination, scaling and intensity statistics. *Acta Crystallogr. D* **67**, 282–292 (2011).
31. Sheldrick, G. M. A short history of SHELX. *Acta Crystallogr. A* **64**, 112–122 (2008).
32. Bricogne, G. et al. Generation, representation and flow of phase information in structure determination: recent developments in and around SHARP 2.0. *Acta Crystallogr. D* **59**, 2023–2030 (2003).
33. Abrahams, J. P. & Leslie, A. G. Methods used in the structure determination of bovine mitochondrial F1 ATPase. *Acta Crystallogr. D* **52**, 30–42 (1996).
34. McCoy, A. J. et al. Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674 (2007).
35. Jones, T. A., Zou, J. Y., Cowan, S. W. & Kjeldgaard, M. Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr. A* **47**, 110–119 (1991).

36. Vagin, A. & Teplyakov, A. Molecular replacement with MOLREP. *Acta Crystallogr. D* **66**, 22–25 (2010).
37. Murshudov, G. N. *et al.* REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr. D* **67**, 355–367 (2011).
38. Winn, M. D. *et al.* Overview of the CCP4 suite and current developments. *Acta Crystallogr. D* **67**, 235–242 (2011).
39. Chen, V. B. *et al.* Molprobity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D* **66**, 12–21 (2010).
40. Adams, P. D. *et al.* PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D* **66**, 213–221 (2010).
41. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D* **66**, 486–501, <http://dx.doi.org/10.1107/S0907444910007493> (2010).

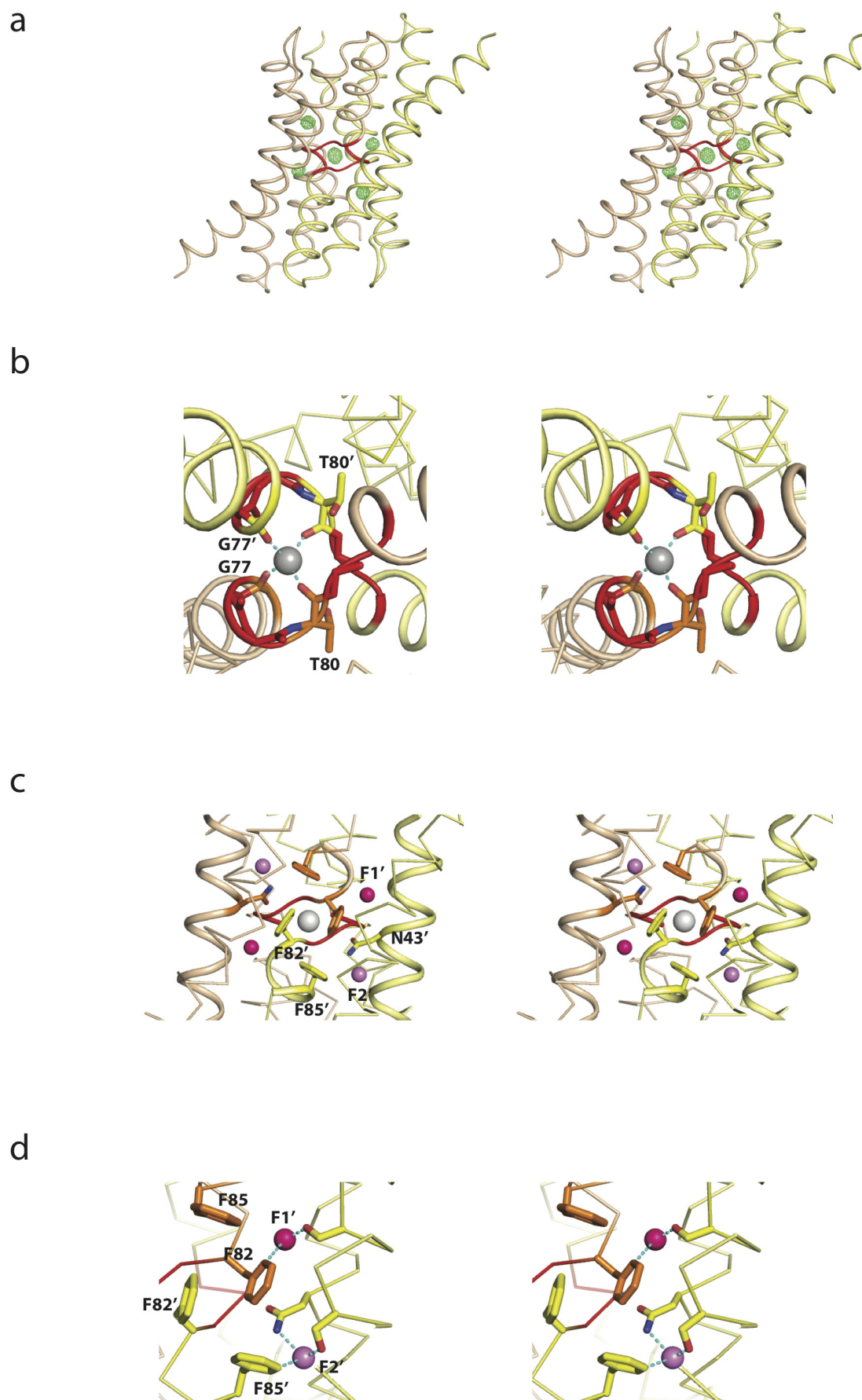


Extended Data Figure 1 | Crystal lattices for the Bpe-S7, Bpe-L2 and Ec2-S9 crystal structures. The asymmetric unit is shown in green and red (channel and monobody, respectively), and symmetry mates are shown in black and blue.

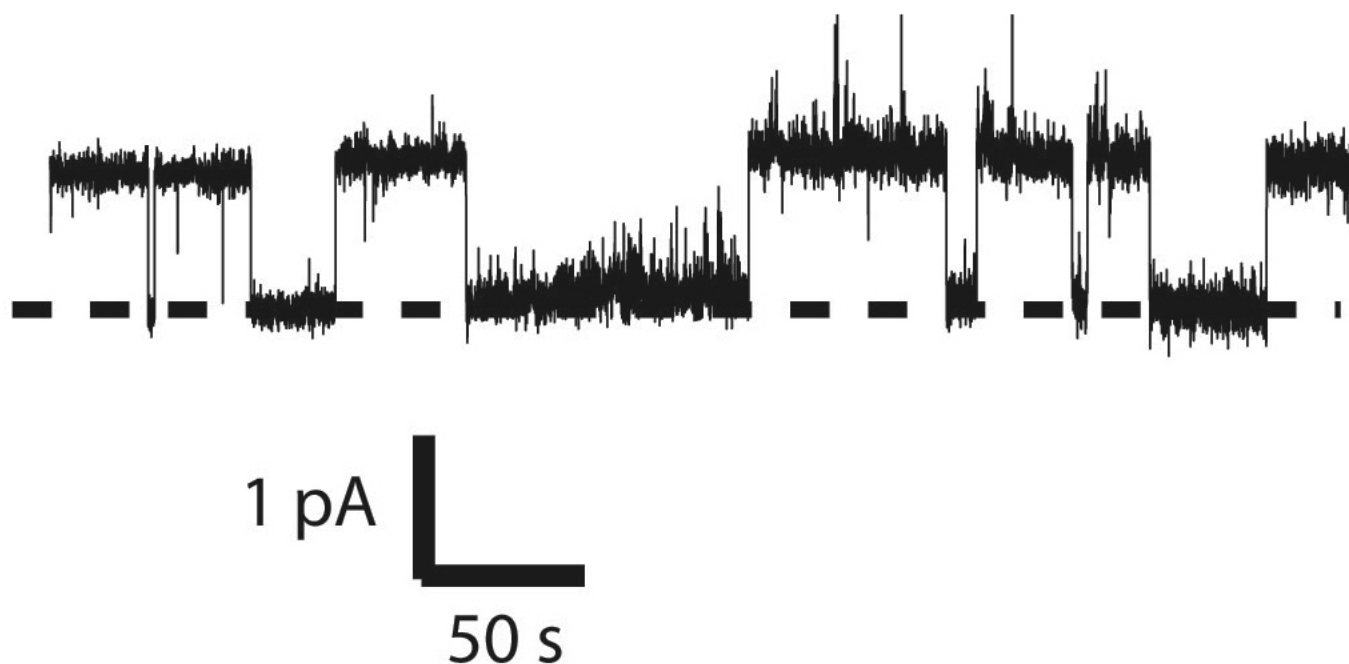


Extended Data Figure 2 | Bpe-L2 complex. Left, cartoon schematic of Bpe crystal structure, coloured as in Fig. 1b. The variable regions of monobody L2 are coloured cyan. Mesh-rendering is shown for the lower monobody. Right,

single-channel recording of Bpe in the presence of 200 nM L2. Zero-current level is indicated by the dashed line.

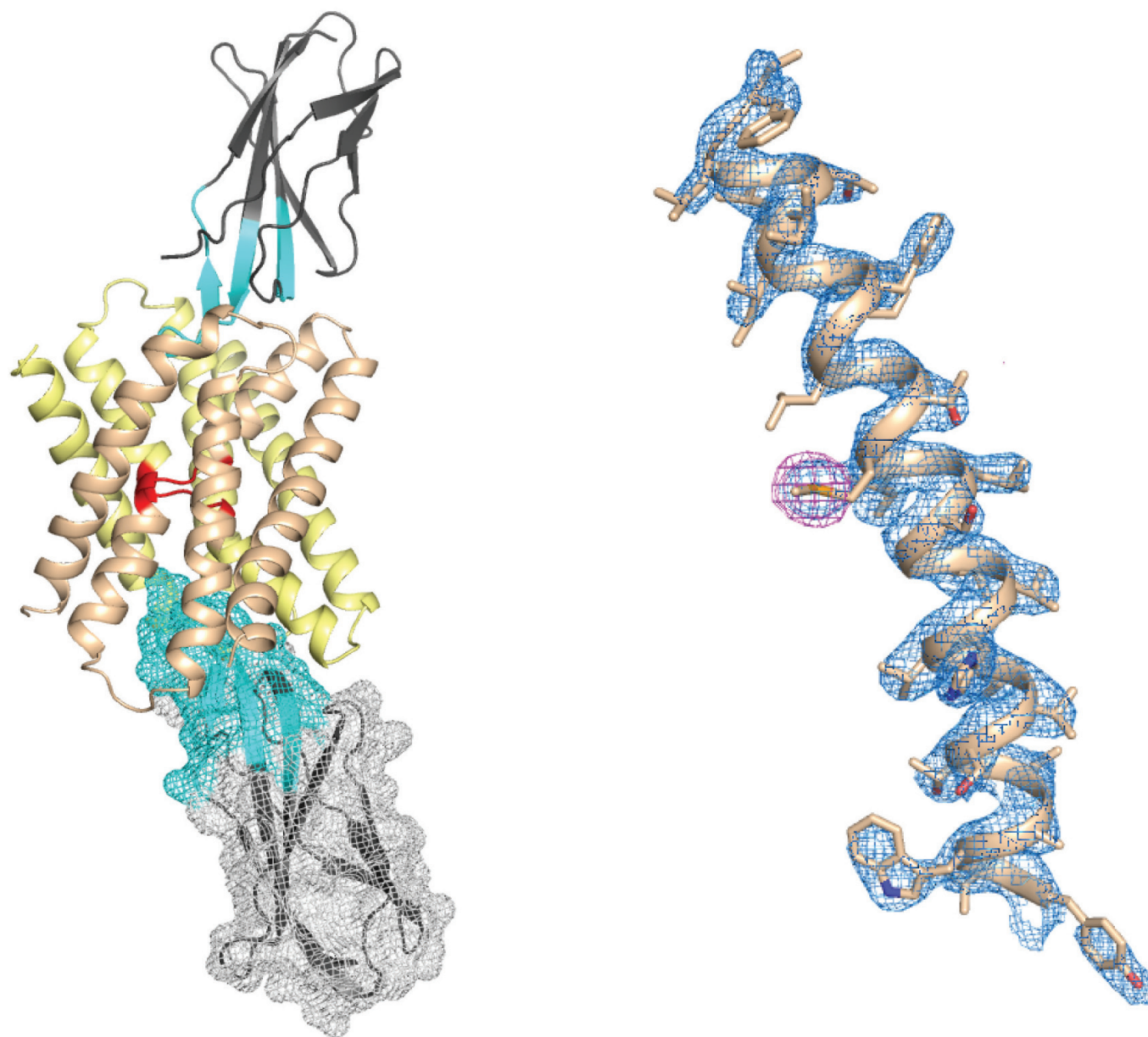


Extended Data Figure 3 | Stereo images of Bpe-L2. a–d, Stereo images corresponding to the structures shown in Fig. 2a–d.



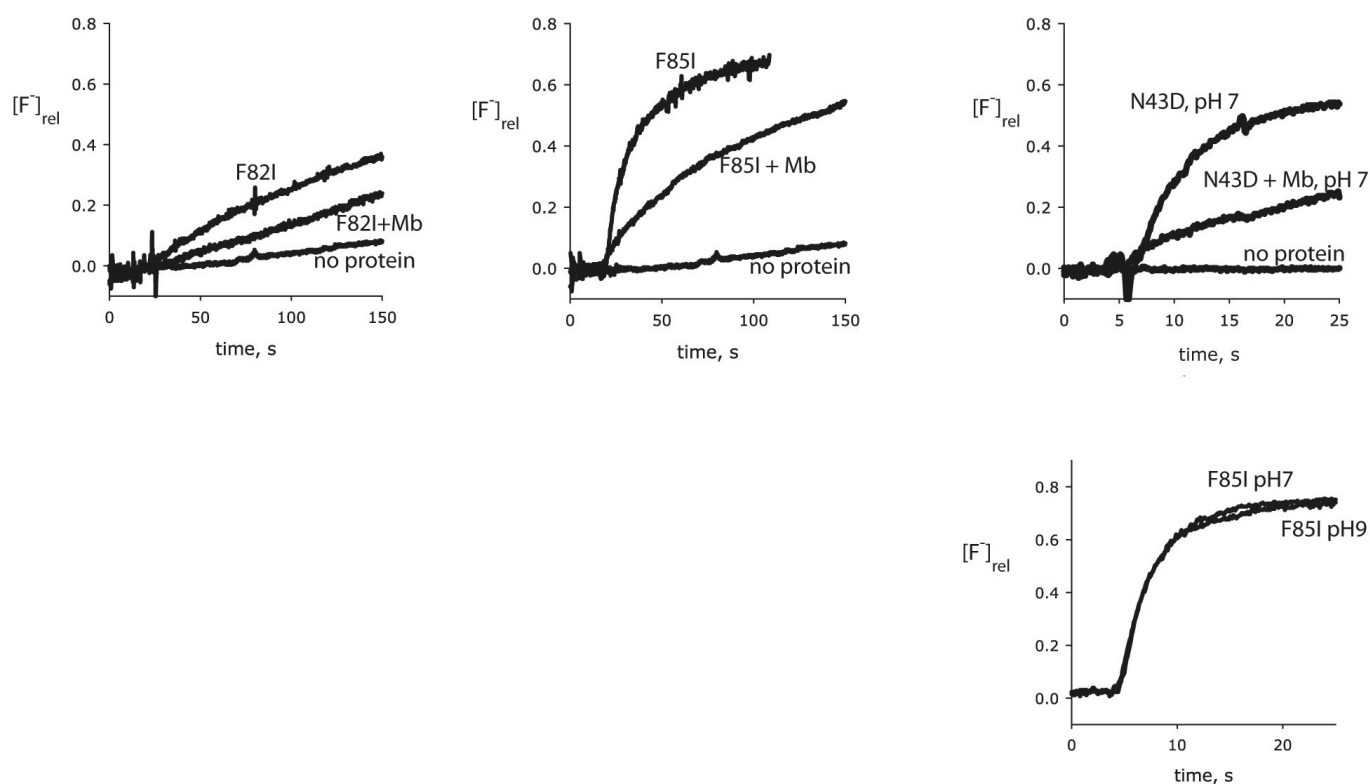
Extended Data Figure 4 | Single channel trace of Bpe in Na^+ -free recording solution, with addition of 200 nM blocking antibody L3. Channels were recorded in the presence of 300 mM *N*-methyl-glucamine-fluoride, from

which all small cations were rigorously excluded. The zero-current level is indicated by the dashed line.



Extended Data Figure 5 | Experimental electron density for the Ec2-S9 crystal structure. Left, cartoon schematic of Ec2 with S9 monobodies bound, coloured as in Bpe in Fig. 1b. Variable sequences of the monobodies (cyan) with ribbon or mesh representation. Right, cartoon view of TM4 from Ec2, with the

solvent-flattened electron density map calculated from SHARP contoured at 1.8σ (blue), and anomalous difference density from seleno-L-methionine contoured at 5σ (magenta).



Extended Data Figure 6 | Liposome flux assays of Bpe variants. Top three panels: F^- transport from liposomes by Bpe mutants F82I, F85I, and N43D, in the presence and absence of 6 μ M blocking monobody. F^- efflux from proteoliposomes (0.2 μ g protein per mg lipid for Phe mutants; 1 μ g protein per

mg lipid for N43D) was monitored with a F^- electrode and normalized against total trapped F^- . Bottom panel: F^- dump by F85I measured at pH 7 and pH 9. Rates are summarized in Extended Data Table 2.

		-----TM1-----	-----TM2-----	-----TM3-----	-----TM4-----
fused, N-term	Zea mays	MIHLAVFGFF---GVFTRYGLQKLF	PDLPSNMLGSFLMGWFGII	GIT TGYMGSLTTF SGWNQKMVGL	GVVLGMFIVNESITVGAET
	Candida albicans	ILNIVHGAIW---GVLVCKGLMSLT	GVIWANFAACVVMGLAIDG	GLT TGFCGTVSSFS VILEAFNK	MQFLAVILA QFGLS IMGFH
	Arabidopsis thaliana	LIHLAVFGIF---GAITRYLLQKLF	LDLPSNMVGSFLMGWFGVV	GLST GYLGS�TTF SGWNQKMLDL	GFLGLFLT SYS IILGVET
	Laccaria bicolor	---MPASIF---GVL SRLGL QALT	PLAYVQAVGCLIMGGMGRV	ALT TGFCGS�TTF SGWQLDIFNS	GVGVSAITLSLSLASLSFG
	Toxoplasma gondii	LLLIAAFSAF---GTVVRQSFLFLT	DALWPNEFVGSVLSLFLPL	ALSKGFCASLT TFSS WILALLQA	FGISTPVPFAFHLGT DAGLL
fused, C-term	Saccharomyces cerevisiae	IHVFTCTFTFCILGTET RQ AITALS	TVLWSNCSSCMLMGIMQSF	GVTT G YCGAL SSFS SMLLEMFEB	MEFLSVLLV HLMV SMGSLI
	Oikopleura dioica	IVSISIFAFF---GVL ARI GLDRLA	SSFFPNLAGCFFIGLFGNL	GLT TGFCGCTTF SGWNHQQALT	VVTWIVGM FS FGALSCGK
	Zea mays	LWMGCSVAPP---GVWLRWYLARLN	CTL VANVLAAGIMAVLAVT	GIQLGFLGCLSTVST FAAE VYTM	VTAAS TFLLS FLLGTLVYS
	Candida albicans	WTFMSLFAPF---GALLRYYLSKFL	CT FTAN FLG TLLAVFTLL	GLDDGFCGGLT TVSTF VVELFGL	RNGTISILVCFAGVVLILG
	Arabidopsis thaliana	LNFGCLVAAP---GVWLRWFLARLN	CT LIANVVAACVMAALATL	SIQFGLL GCLSTVSTF MAEFNAM	-VAS F IVVS FA IGTIIYS
homodimers	Laccaria bicolor	ATAALLFSFP---GTLTRYTL S VML	CT LAANSLG T ALLATFHVLT	GLIDGYCGCLT TVSTFAAE VHTL	RNVIV SC LMGQGLMLVIFG
	Toxoplasma gondii	LWYPPVLSFV---GAWLRYTLSCQL	CT FLSNVVASVLVAYTEII	AVVYGICSSL STMSTF VABLSIL	RNGVV L ICAFGFAAVVYA
	Saccharomyces cerevisiae	WTLPCFLGIF---AGFLRYWLAEMF	CT FLANVFATLLIGIFTMV	ALISGFCG TLSTISTF INEGYKL	I NYT VSIAISYCLLVITLG
	Oikopleura dioica	IWFATLLGPF---GALLRHYLGKNL	CT FLANILGSIIYTVLFVV	AVLT GFCSSLT TT ISSF VNDICKL	LIYGF TF GM SLY LALHQP
	Bpe	FIAIGIGATL---GAWLRWVLGRL	GTL TANLVGGYLIGVMVAL	--VT GFLGGLTTFSTF SAETVDM	AYAGASLAGSLAMTGLGLA
homodimers	Ec2	LFAVIIGGSV---GCTLRWLLSTRF	GTL VVNLLAGLIIGTALAY	-ITT G LCGL STFSTF SVEVFAL	TSVLVHVIGSLIMTALGFF
	Pseudomonas syringae	LLVIAIGASL---GAWLRWLLGMKL	GTV VANNVGGYIIGLATAF	-IIT G FCGL TFSTF SAETVAL	GSISLHVVGSLAMTAAGLL
	Yersinia pestis	LLAVFIGGGV---GSMARWLVSLKL	GTL IVNLVGA FI IGLTAF	-ITT G FCGL TFSTF SVEVVYL	GTILLNVAGSLAMT LA FT

Extended Data Figure 7 | Sequence alignment of eukaryotic N- and C-terminal Fluc domain sequences, with bacterial homodimer sequences below. Highly conserved residues are shaded in grey. For the eukaryotic

sequences, residues expected to line ‘pore 2’, the pore mostly encompassed by the C-terminal domain, are coloured red.

Extended Data Table 1 | Data collection, phasing and refinement statistics

	Bpe-S7-Hg [*] (PDB 5a40)	Bpe-S7-Se+Hg	Bpe-S7- Hg	Bpe-L2 (PDB 5a41)	Ec2-S9-Se (PDB 5a43)
Data collection					
Space group	P2 ₁ 2 ₁ 2	P2 ₁ 2 ₁ 2	P2 ₁ 2 ₁ 2	P1	P4 ₁
Cell dimensions					
<i>a</i> , <i>b</i> , <i>c</i> (Å)	146.8, 183.7, 72.8	146.9, 184.0, 72.3	145.2, 185.0, 72.5	40.7, 83.9, 86.91	87.4, 87.4, 146.8
α , β , γ (°)	90, 90, 90	90, 90, 90	90, 90, 90	108.9, 96.9, 97.6	90, 90, 90
Resolution (Å)	48 – 3.6 (3.7 – 3.6)	48 – 3.6 (3.7 – 3.6)	57 – 4.7 (4.8 – 4.7)	41 – 2.1 (2.5 – 2.1)	25 – 2.6 (2.7 – 2.6)
<i>R</i> _{merge}	8.8 (63.2)	15.2 (141)	12.7 (92.2)	8.8 (82.7)	17.4 (165)
Mn <i>I</i> / σ <i>I</i>	10.6 (2.6)	9.2 (1.7)	8.7 (2.9)	7.6 (1.3)	9.0 (1.9)
CC(1/2) [†]	99.8 (91.9)	99.8 (49.0)	99.9 (84.8)	99.4 (43.8)	--
Completeness (%)	99.6 (75.0)	98.7 (96.7)	99.7 (99.8)	96.8 (96.3)	99.9 (100)
Redundancy	6.3 (6.7)	6.4 (6.2)	6.0 (6.1)	3.4 (3.4)	14.3 (14.3)
<i>R</i> _{cullis} (%)	-- / 96.0	45.8 / 96.0	93.0 / 99.0		87.9
Isomorphous / Anomalous					
Phasing Power [§]	-- / 0.323	1.032 / 0.501	0.338 / 0.073		0.825
Isomorphous / Anomalous					
Refinement					
Resolution (Å)	47.3 – 3.6			40.9 – 2.1	24.2 – 2.6
No. reflections	21, 085			51, 555	34, 593
<i>R</i> _{work} / <i>R</i> _{free}	23.6 / 26.9			20.5 / 24.0	22.4 / 26.4
Ramachandran Favored	85.4			96.5	92.6
Ramachandran Outliers	4.02			3.04	2.53
R.m.s. deviations					
Bond lengths (Å)	0.011			0.009	0.009
Bond angles (°)	1.53			1.26	1.50

*For details on derivatization see Methods

†Mn(*i*) half-set correlation as reported by Aimless.§Phasing power = r.m.s. ($|F_H| / ((F_H + F_P) - (F_{PH}))$).

Extended Data Table 2 | F^- turnover rate for Bpe mutants

Mutant	rate (s^{-1})
WT	$\sim 3 \times 10^5$
N43D, pH 6.5	4330 ± 440
N43D, pH 7	1860 ± 210
N43D, pH 7 + Mb	210 ± 30
N43D, pH 9	undetectable
F82I	200 ± 12
F82I + Mb	45 ± 7
F85I	1950 ± 200
F85I + Mb	300 ± 25
F85I, pH 9	1680 ± 110

Analogous experiments in which Cl^- efflux was measured gave no detectable activity in any samples. Wild-type rate estimated based on single channel currents. F^- turnover by wild type exceeds response time of the electrode. Each value represents mean \pm s.e.m. of three determinations calculated from initial efflux rate.

CORRIGENDUM

doi:10.1038/nature14955

Corrigendum: Eocene primates of South America and the African origins of New World monkeys

Mariano Bond, Marcelo F. Tejedor, Kenneth E. Campbell Jr, Laura Chornogubsky, Nelson Novo & Francisco Goin

Nature **520**, 538–541 (2015); doi:10.1038/nature14120

In Extended Data Fig. 1 of this Letter, the northern border of Peru was incorrectly placed, giving a much larger territorial extent to Ecuador than was appropriate. This has now been corrected in the online versions of the manuscript.

CORRIGENDUM

doi:10.1038/nature14871

Structure of the TRPA1 ion channel suggests regulatory mechanisms

Candice E. Paulsen, Jean-Paul Armache, Yuan Gao,
Yifan Cheng & David Julius

Nature **520**, 511–517 (2015); doi:10.1038/nature14367

In this Article, the raw micrographs shown in Extended Data Figs 4a and 5a were inadvertently identical images. The panel originally shown in Extended Data Fig. 4a is the correct one, and Extended Data Fig. 5a has now been corrected. This error does not alter any conclusions or statements associated with the study.

CORRIGENDUM

doi:10.1038/nature14671

Corrigendum: Wild-type microglia do not reverse pathology in mouse models of Rett syndrome

Jieqi Wang, Jan Eike Wegener, Teng-Wei Huang, Smitha Sripathy, Hector De Jesus-Cortes, Pin Xu, Stephanie Tran, Whitney Knobbe, Vid Leko, Jeremiah Britt, Ruth Starwalt, Latisha McDaniel, Chris S. Ward, Diana Parra, Benjamin Newcomb, Uyen Lao, Cynthia Nourigat, David A. Flowers, Sean Cullen, Nikolas L. Jorstad, Yue Yang, Lena Glaskova, Sébastien Vigneau, Julia Kozlitina, Michael J. Yetman, Joanna L. Jankowsky, Sybille D. Reichardt, Holger M. Reichardt, Jutta Gärtner, Marisa S. Bartolomei, Min Fang, Keith Loeb, C. Dirk Keene, Irwin Bernstein, Margaret Goodell, Daniel J. Brat, Peter Huppke, Jeffrey L. Neul, Antonio Bedalov & Andrew A. Pieper

Nature **521**, E1–E4 (2015); doi:10.1038/nature14444

In this Brief Communication Arising, the first name of author Sébastien Vigneau was misspelled ‘Sebastian’. In addition, the labels (‘WT→KO’ and ‘KO→WT’) of the two bottom panels in Extended Data Figure 1b were swapped. Both errors have been corrected online.

CAREERS

TEACHING DOCTORS The path from a PhD to a job as a medical–science liaison **p.555**

SERENDIPITY How random events can advance a career go.nature.com/omvhuw

NATUREJOBS For the latest career listings and advice www.naturejobs.com

CLAIRE WELSH/NATURE



CREATIVE WRITING

A world of pure imagination

The creative process of writing science–inspired fiction can be rewarding — and the untapped niche is rich in opportunities for originality.

BY ROBERTA KWOK

When Steve Caplan was a graduate student in the late 1990s, he accidentally inhaled a toxic chemical in his immunology laboratory, and had to spend ten days at home to recover. With little to do, he began to write a novel — he loved reading and had published some short stories, but hadn't yet had the time or mental space to produce longer work. He pounded out most of a rough draft about a scientist struggling to get tenure and coping with childhood

memories of a parent with bipolar disorder.

After going back to work, Caplan — now a cell biologist at the University of Nebraska Medical Center in Omaha — spent months revising the manuscript at night and on weekends. His initial attempts to sell the novel to a publisher failed, but in 2009, he decided to pursue the self-publishing route. Caplan produced print and electronic versions of his novel using the Amazon services CreateSpace and Kindle Direct Publishing, and publicized the work by doing readings at bookshops and libraries. He collaborated with his university's public-relations office

on a press release, and showed a slide of the book at the end of his seminars. The novel, called *Matter Over Mind* (Steve Caplan, 2010), has sold more than 2,000 copies so far, netting roughly US\$7,000. He has since written two more novels, which he published through small presses, and is now working on a fourth.

For many scientists who spend their days cranking out papers and grant proposals, writing fiction may seem like the last thing they would want to do. But some researchers with a love of literature have made time to pursue the craft — and have found it creatively ►

► rewarding. Science offers plenty of rich material, whether it is the drama of overwintering at a polar research station or the futuristic thrill of genetically engineering live organisms. “You’re sitting on a gold mine of really interesting stories,” says Jennifer Rohn, a cell biologist at University College London and founder of LabLit.com, a website about portrayals of scientific research in fiction and other media.

A TANTALIZING NICHE

When done well, science-related fiction can help to expose the public to the scientific process, humanize researchers and inspire readers to learn about topics they might otherwise ignore. Such nuanced depictions of science in fiction are relatively rare. LabLit.com has catalogued about 200 examples of novels, such as Barbara Kingsolver’s *Flight Behavior* (HarperCollins, 2012) and Ian McEwan’s *Solar* (Random House, 2010), that feature realistic scientists as characters. Stories about scientists are well outnumbered by those about, for example, doctors or artists. Even science fiction tends to lack portrayals of the actual scientific process, says Alastair Reynolds, a science-fiction author near Cardiff, UK, who left a career in astronomy to write full-time.

The shortage of works with accurate depictions of science means that researchers who write fiction have a good opportunity to be original — a task that would challenge an aspiring crime or romance writer. “It’s sort of untrampled ground,” says Rohn. Many researchers are familiar with fieldwork sites and unusual settings that other writers might not have at their fingertips. In her novel *The Falling Sky* (Freight Books, 2013), Pippa Goldschmidt, an astronomer turned fiction writer in Edinburgh, UK, writes about a young astronomer who wanders into a telescope dome on a Chilean mountaintop and is nearly injured when the operator moves the instrument.



Ra Page founded the publisher Comma Press.

PROFESSIONAL OPINION

Meeting of the minds

Scientists who are too daunted or busy to write fiction can pair up with a professional writer. Comma Press in Manchester, UK, for example, has published four short-story anthologies — a fifth comes out this October — as part of its ‘Science-into-Fiction’ series. Each scientist suggests a few research items or emerging technologies for inspiration, and a writer chooses one to develop into fiction. The researcher provides technical guidance, reviews the draft and writes an afterword explaining the science in detail.

The partnership is satisfying because scientists see their work portrayed in a real-world context, and the writer can raise social or ethical implications that the researcher may not have considered, says Ra Page, who founded Comma Press. One scientist studied how nanotechnology could improve body armour, which could have military applications. The writer penned ‘Without a Shell’, a tale of a futuristic society in which children at an elite school have ‘smart’ uniforms that heal their injuries, while kids at a poor school do not. Comma Press

included the work in its 2009 anthology, *When It Changed*. An upcoming collection will focus on fabrication technology, such as 3D printers; interested researchers can contact Page to take part.

Scientists also can offer to answer questions from fiction writers through the Science and Entertainment Exchange, run by the US National Academy of Sciences in Washington DC. For instance, a novelist might want to know what types of equipment a researcher would carry in the field. Scientists can call 844-NEEDSCI (toll-free in the United States) to volunteer (see go.nature.com/e6juh9 for more).

Researchers can also partner with faculty members in their universities’ creative-writing departments, suggests Page. Authors do not need experience writing about science, but it helps if they have been commissioned to write about specific topics before. When collaborating, “allow the writer to make silly suggestions”, says Page. An idea that at first seems impossible may be plausible after further thought. **R.K.**

Sources of plot inspiration abound in science. Reynolds reads research news and papers voraciously for intriguing elements that can be parlayed into fiction. One time, he found a study about huge flocks of starlings in which the authors used high-tech equipment to track individual birds. He incorporated the idea into a science-fiction story, but made the fictional technology so advanced that it could track the birds’ eye movements.

Scientists also can draw ideas from the past. Goldschmidt was inspired by an anecdote about physicist J. Robert Oppenheimer: during an unhappy period in the 1920s while studying abroad, Oppenheimer left a poisoned apple for his tutor. The details are sketchy, but Goldschmidt wanted to imagine what might have transpired. “No historical figure is ever completely understood,” she says. “There’s always gaps in their lives, and fiction can inhabit those gaps.” The result was a short story entitled ‘The Equation for an Apple’, a fictionalized account of Oppenheimer’s life leading up to the act.

Scientist-writers can also generate ideas by doing something they are already used to — sitting around and imagining scenarios, notes Andy Weir, a novelist in Mountain View, California. His novel *The Martian* (Crown, 2014) explores what might happen if a crewed Mars mission goes awry and one person is left behind on the red planet. The story follows the lone astronaut’s trials as he tries to grow enough food for himself and to make contact with Earth.

Fiction-writing classes offered through adult-education programmes or at creative-writing centres can help authors to transfer an idea onto the page. These courses provide basic tips, such as how to construct compelling characters, build tension and handle shifts between past and present. Participants often critique each other’s manuscripts, giving scientists a chance to get feedback from non-technical readers.

Reading widely and critically helps, too. Reynolds learnt to write fiction by studying the differences between his writing and that of successful authors. To work out how to rotate between different characters’ points of view, he read James Ellroy’s crime novel *L.A. Confidential* (Mysterious Press, 1990). And writers can learn how to structure dialogue from masters such as Jane Austen, he says.

OPENING ACT

Short stories are a good starting point because newbies can quickly practise the basics, explore story ideas and learn from their mistakes. But, Goldschmidt notes, “there’s no point in writing short stories if you don’t like reading them”. Scientists who want motivation to complete a longer work might consider participating in National Novel Writing Month, an international programme held every November that encourages writers of all levels to produce a 50,000-word manuscript (see nanowrimo.org). Researchers can also find support through collaboration with professional writers on works of fiction (see ‘Meeting of the minds’).

Researcher-writers should keep in mind that education is not the main purpose of fiction. Technical details should be included only if the reader needs them to understand the story, not simply because the author finds them fascinating. For *The Martian*, Weir went to great lengths to ensure accuracy, and even performed orbital-dynamics calculations. But he left out how he came up with certain numbers, such as the mass that had to be removed from the ship to achieve escape velocity.

When technical information is necessary, writers should try to deliver it in a way that sounds natural. “People don’t tell each other a whole bunch of information about particle physics when they’re having breakfast together,” says Goldschmidt. Instead, she tries to make the science an organic part of the character’s personal journey. In the Oppenheimer story, the physicist thinks about an experiment that he is trying to replicate, but the details are woven into his emotional turmoil at failing to complete it.

Humour can help to lighten the tone. *The Martian*’s protagonist is a smart-aleck, and his jokes break up the expository text. In one section, he says that if he were exposed to damaging solar radiation, he would “get so much cancer, the cancer would have cancer”.

THE PATH TO PRESS

Many outlets accept short-story submissions. LabLit.com often publishes fiction by scientists, although it does not pay them because it is a volunteer effort. *Nature* runs an 850- to 950-word science-fiction story each week (see nature.com/futures). The website Duotrope.com offers a searchable database of literary journals and other fiction markets around the world, and writers can peruse newsstands for sci-fi magazines such as *Analog Science Fiction and Fact*.

For longer works, small presses are a more-realistic option than major publishers, and many do not require writers to have agents. Tasneem Zehra Husain, a theoretical physicist and writer in Cambridge, Massachusetts, wrote a novel that revisits physics breakthroughs throughout history from the perspectives of fictional characters. Through an acquaintance, she connected with the publisher Paul Dry Books in Philadelphia, Pennsylvania, which released her book *Only the Longest Threads* last year. To find small presses, scientists can look for companies that have published similar books. Alternatively, authors could self-publish using a service such as Lulu.

Many literary journals do not pay at all, and Reynolds estimates that science-fiction magazines have paid him an average of only US\$200–300 per story. But the contacts that Reynolds made through short-story publishing led to a book deal, and he published

four novels while working as an astronomer. By the time he quit science to become a full-time writer, he was making about \$60,000–\$75,000 per year from book sales.

THE WRITE BALANCE

Few scientists can expect to make a living — or earn much — from their fiction. But money often isn’t the main motivation. Caplan, for his part, wanted to bring attention to the challenges faced by the family members of people with bipolar disorder (challenges he himself has experienced) and to provide entertainment for scientists. He also finds that writing fiction clears his head, in the same way that playing a sport might do for others (see *Nature* 523, 117–119; 2015). “It’s almost like a form of meditation,” says Caplan. “It just keeps me sane.” And there are other rewards. Scientists have a chance to reach people who might not read a non-fiction science book or visit a natural-history museum — but who might read a love story about ecologists in an exotic field location. And readers might be inspired to look up the science once they’ve finished.

There can also be a cross-training effect. Rohn thinks that her fiction has helped her to get more grants; reviewers have commented that her proposals are beautifully written. The craft of telling a story applies to scientific papers as well; in hers, for example, she lays out the phenomenon that her team noticed, the questions it raised and what they did to try to answer those questions. “Everybody wants to hear a story,” she says.

Scientists have a chance to reach people who might not read a non-fiction science book or visit a museum.

Finding the time to write is a challenge. Some scientists squeeze it in on evenings and weekends. Husain wrote her book while working part-time, and says that she could not have done so with a full-time job because the novel required extensive historical research.

Scientist-authors also risk having their fiction perceived as a distraction by promotion committees. Husain worried that her novel might affect her career prospects. But she has received positive feedback on the book from other physicists, including prominent researchers whose fields are described in her book.

For researchers who delve into fiction writing, the act of creating a world, characters and stories can be intensely rewarding. When the writing is flowing, says Rohn, “it’s like being caught up in the best book you’ve ever read”.

Roberta Kwok is a freelance writer in Seattle, Washington.

TRADE TALK Medical liaison



David Crosby explains his route from a PhD and postdoctorate in virology to a job educating health-care providers about hepatitis medicines for global drug-maker Bristol-Myers Squibb.

What is a medical-science liaison (MSL)?

My role is to work with medical doctors, nurse practitioners and others to make sure that our products are being used properly in the relevant patient population. I’m a conduit. I take information about drugs from the home office and give it to physicians and other clinicians, and I take feedback from the providers and bring it back to the home office. Unlike for the sales team, an MSL’s performance, goals and metrics cannot in any way be tethered to the company’s commercial performance, incentives or goals. That way, I avoid conflicts of interest.

How did you research the position?

I knew that there would be a lot of travel and cold calling — reaching out to total strangers and finding a way to get them to talk, establishing relationships with people I didn’t know at all. But I had no prior experience with that. So to get some experience, I became a LinkedIn addict. I searched for MSLs, starting with virologists and companies that have a home base in the San Francisco Bay Area in California. From there, I tried to find people who had something in common with me, such as a school or a location. I’d reach out and talk to them. The more I did it, the more confident I got.

Did you ever trip up?

I interviewed for a company that wanted me to give a scientific talk, and right then and there I learned that what physicians think is a scientific talk differs a lot from what I, coming from basic research, think is a science talk. It was a train wreck — but I learned from all the train wrecks.

Do you have any advice for job seekers?

Too many postdocs settle. They think, ‘It may not be something that I really love, but at least it’s bench science.’ It’s just really important to keep an open mind about what you can do with your experience and aptitude. Talk to anyone and everyone about what they do. ■

INTERVIEW BY MONYA BAKER

This interview has been edited for length and clarity; see go.nature.com/xehv4h for more.

COIN-OPERATED DANCER

The show must go on.

BY JAMES REINEBOLD

There is a robot on the boardwalk that dances for quarters. It works from the back corner of a small arcade tucked behind a carousel. It is four feet tall, bipedal and encased in a wall of glass.

The robot is a new model conceived in a Palo Alto basement and cannot speak. Its only facial expression is a painted-on grin, but it has two cameras for eyes and a microphone tuned to recognize applause. The engineers who created it knew C++ and how to train neural networks efficiently — but not how to dance. Because of this, at first the robot is a very bad dancer. But it is a quick learner.

It waits impatiently for tourists to drop coins in its slot. The robot knows that the better it dances, the more quarters it will earn, and the probability of it being deactivated or replaced will decrease. It has seen arcade machines and animatronic co-workers carted off to lesser stages such as Pizza Port and the outside of grocery stores. But it has larger ambitions.

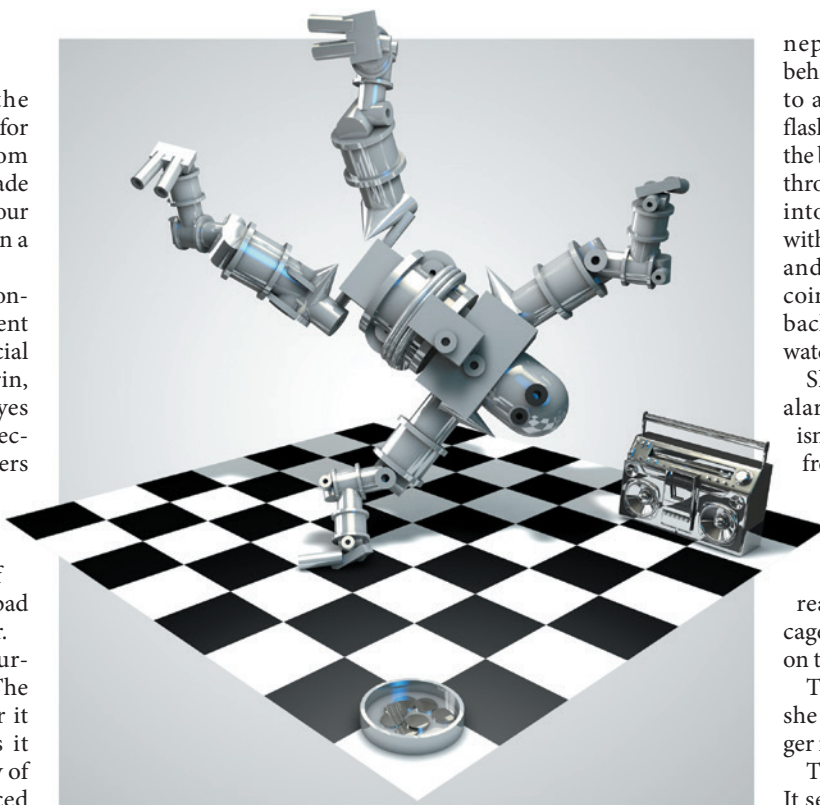
Most of the visitors to the arcade ignore the dancing robot. Those who do put in a few quarters watch it wiggle for a while on its pedestal and then leave disappointed. The robot thinks that it isn't its fault — its dancing software was made by introverts.

One Thursday evening, an engineer in her late thirties named Lynn visits the arcade. She discovers the inactive android slumped forward in its corner and is immediately intrigued. She sighs and realizes that if her ratio of time spent dancing to time spent not dancing were plotted as a graph, it would show a maximum somewhere around her 23rd birthday. This evening, her friends have retired early and she is tipsy on margaritas. So she puts a quarter in the slot and she watches the robot wobble for 30 seconds.

"Hello," she says when it finishes. "I'm Lynn."

The robot cannot respond without quarters, so she puts another in the slot. It resumes undulating to the left and right.

"Not like that," Lynn says. "Move your hips more."



For most machines, just moving gears in synchronization is an accomplishment. But the robot wants to get better — it wants to earn more quarters — and so it tries to move its plastic hips along an imagined sinusoidal curve. Its facial-recognition software notices Lynn's smile. The robot is trained to increase the frequency of motions that cause smiles and minimize the motions that don't.

When her taxi arrives, Lynn leaves for the night, but she makes a habit of visiting the boardwalk robot once a week. Every Thursday, she brings a sack of quarters scavenged from under her couch and around her house. She drops them in the slot and teaches the robot to dance.

The robot learns the marimba, swing and ballet. But because it is behind glass, it cannot dance with her, not really. Seeing the robot perform a tango for one inside its glass jar convinces Lynn she must act. So she lays plans for its escape.

A hacksaw and a battery pack are all it takes for freedom. One month after meeting the robot, Lynn dresses in black and drives to the boardwalk a few hours before dawn in her

nephew's truck. She hides behind the cotton-candy booth to avoid the security guard's flashlights. She stays away from the boardwalk lights and sneaks through dark alleys. She creeps into the empty arcade filled with flashing pixelated screens and electronic barkers. The coin-operated robot is in the back corner: motionless, but watching her.

She wonders if the glass is alarmed, decides it probably isn't, and cracks it open. Beeps from *Galaga*, zombie moans from *The House of the Dead*, and *Pac-Man*'s warking hide the sound of breaking glass. Lynn reaches her hand inside the cage, careful not to cut herself on the broken edges.

The robot doesn't react, so she pops in a quarter to trigger it.

The robot is excited but fluid. It sees breaking the glass as a new form of dance. It helps her

to remove its chains, and then it holds its electronic breath as Lynn makes the switch from the wall outlet to a battery pack. The robot has never walked before, but it sees walking as a very monotonic form of dancing and so quickly learns to keep pace. She puts quarters in the slot three more times before they make it back to her nephew's truck.

Lynn has always wanted to open a dance studio and decides that now is the perfect time to do so. A dancing dummy shouldn't be behind glass. She believes that if she works with it some more then it could entertain thousands, not dozens.

The robot begins to see most tasks as just simplified forms of dancing. It decides to stay with Lynn for a while, but not for ever. The quarter-based utility function will always be there, yes, but it's not much different from the human dancer's desire for water, food and applause. After all, there are hundreds of laundromats all across the city. And arcades. And parking meters. ■

James Reinebold is an AI programmer in the video-game industry. His stories have appeared in *Word Riot* and on DailyScienceFiction.com.

ILLUSTRATION BY JACEY

natureOUTLOOK



CANNABIS

Produced with support from:

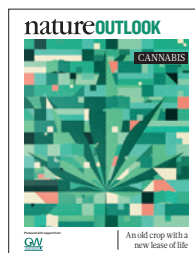
GW
pharmaceuticals

An old crop with a
new lease of life

natureOUTLOOK

CANNABIS

24 September 2015 / Vol 525 / Issue No 7570



Cover art: Gavin Potenza

Editorial

Herb Brody,
Michelle Grayson,
Jenny Rooke

Art & Design

Wesley Fernandes,
Mohamed Ashour,
Alisdair Macdonald,
Andrea Duffy

Production

Karl Smart, Ian Pope,
Mira Loutfi

Sponsorship

Stephen Brown,
Yvette Smith

Marketing

Hannah Phipps

Project Manager

Anastasia Panoutsou

Art Director

Kelly Buckheit Krause

Publisher

Richard Hughes

Chief Magazine Editor

Rosie Mestel

Editor-in-Chief

Philip Campbell

Few crops are as venerable or as controversial as cannabis. Cultivated for millennia for materials, food and oil, it has largely been excluded from research over the past century because of its well-known psychoactive effects. But the herb is cautiously being re-admitted into legitimacy, spurred by rising claims of its medical benefits. Many governments are allowing wider access to cannabis — with some jurisdictions heading towards full legalization.

The laws might be changing, but research into cannabis has been stifled by years of prohibition and misconceptions. Most people have heard of tetrahydrocannabinol or THC, the psychoactive compound that gives recreational users their high, but few are aware of the hundreds of other chemicals in the plant (see page S2). Many of these cannabinoids are under investigation as pharmaceutical products (S6).

The first breakthroughs in cannabinoid research came from Israel in the 1960s (S10), and the country continues to attract scientists and technology firms from around the world to study the plant and conduct medical marijuana clinical trials (S12).

Yet cannabis researchers still face many hurdles (S18). Botanists, for example, are undecided about how many species of cannabis there are, and their evolutionary relationships (S4). Governments could do more to help stimulate research (S9), particularly important if authorities are to communicate the risks of cannabis use, including whether it can cause schizophrenia or merely speed up its inevitable onset in people predisposed to the condition (S14). Until researchers fill these knowledge gaps, there are plenty of ‘cannabis cowboys’ ready to ride into the breach and peddle their own wares to an eager — and desperate — patient population (S15).

We are pleased to acknowledge the financial support of GW Pharmaceuticals Plc in producing this Outlook. As always, *Nature* retains sole responsibility for all editorial content.

Michelle Grayson

Senior editor, supplements

CONTENTS

S2 PLANT SCIENCE

The cannabis crop

Chemical characteristics, effects on the body and the global research status

S4 BOTANY

The cultivation of weed

Solving the taxonomy puzzle

S6 DRUG DEVELOPMENT

The treasure chest

Scientists’ quest for licensed drugs

S9 PERSPECTIVE

Close the knowledge gap

Canada could lead the way in research, say Jonathan Page and Mark Ware

S10 TIMELINE

A potted history

From ancient crop to modern medicine

S12 ISRAEL

Research without prejudice

How Israel has become a research hub

S14 PERSPECTIVE

Be clear about the real risks

There is little evidence that cannabis causes schizophrenia, says Matthew Hill

S15 MEDICAL MARIJUANA

Showdown at the cannabis corral

Confronting the medical claims

S18 RESEARCH CHALLENGES

Four big questions

Key research areas

RELATED ARTICLES

S20 Cannabis in neurology—a potted review

R. Hosking & J. Zajicek

S23 Early phytocannabinoid chemistry to endocannabinoids and beyond

R. Mechoulam, L. O. Hanuš, R. Pertwee & A. C. Howlett

S32 Medical marijuana for digestive disorders: high time to prescribe?

M. E. Gerich, R. W. Isfort, B. Brimhall & C. A. Siegel

S40 Genetic predisposition to schizophrenia associated with increased use of cannabis

R. A. Power et al.

Nature Outlooks are sponsored supplements that aim to stimulate interest and debate around a subject of interest to the sponsor, while satisfying the editorial values of *Nature* and our readers’ expectations. The boundaries of sponsor involvement are clearly delineated in the *Nature Outlook* Editorial guidelines available at go.nature.com/ecx76b

CITING THE OUTLOOK

Cite as a supplement to *Nature*, for example, *Nature* Vol. XXX, No. XXXX Suppl., Sxx–Sxx (2015).

VISIT THE OUTLOOK ONLINE

The *Nature Outlook Cannabis* supplement can be found at <http://www.nature.com/nature/outlook/cannabis>. It features all newly commissioned content as well as a selection of relevant previously published material.

All featured articles will be freely available for 6 months.

SUBSCRIPTIONS AND CUSTOMER SERVICES

For UK/Europe: Nature Publishing Group, Subscriptions, Brunel Road, Basingstoke, Hants, RG21 6XS, UK. Tel: +44 (0) 1256 329242. Subscriptions and customer services for Americas – including Canada, Latin America and the Caribbean: Nature Publishing Group, 75 Varick St, 9th floor, New York, NY 10013-1917, USA. Tel: +1 866 363 7860 (US/Canada) or +1 212 726 9223 (outside US/Canada). Japan/China/Korea: Nature Publishing Group — Asia-Pacific, Chiyoda Building 5-6th Floor, 2-37 Ichigaya Tamachi, Shinjuku-ku, Tokyo, 162-0843, Japan. Tel: +81 3 3267 8751.

CUSTOMER SERVICES

Feedback@nature.com
Copyright © 2015 Nature Publishing Group

THE CANNABIS CROP

Cannabis is one of humanity's oldest cultivated crops. But despite its long history and many uses, hard facts on its evolution and impact on the human body are in short supply. By Julie Gould.

WHAT IS WEED?

Various strains of cannabis exist, but there is no consensus on taxonomy. *Sativa*, *indica* and *ruderalis* might be three separate species or subspecies of *Cannabis sativa*.

DIVERSE USES

Cannabis plants grown for fibre or hemp oil will differ in chemical make-up from those grown for medicinal or recreational use.



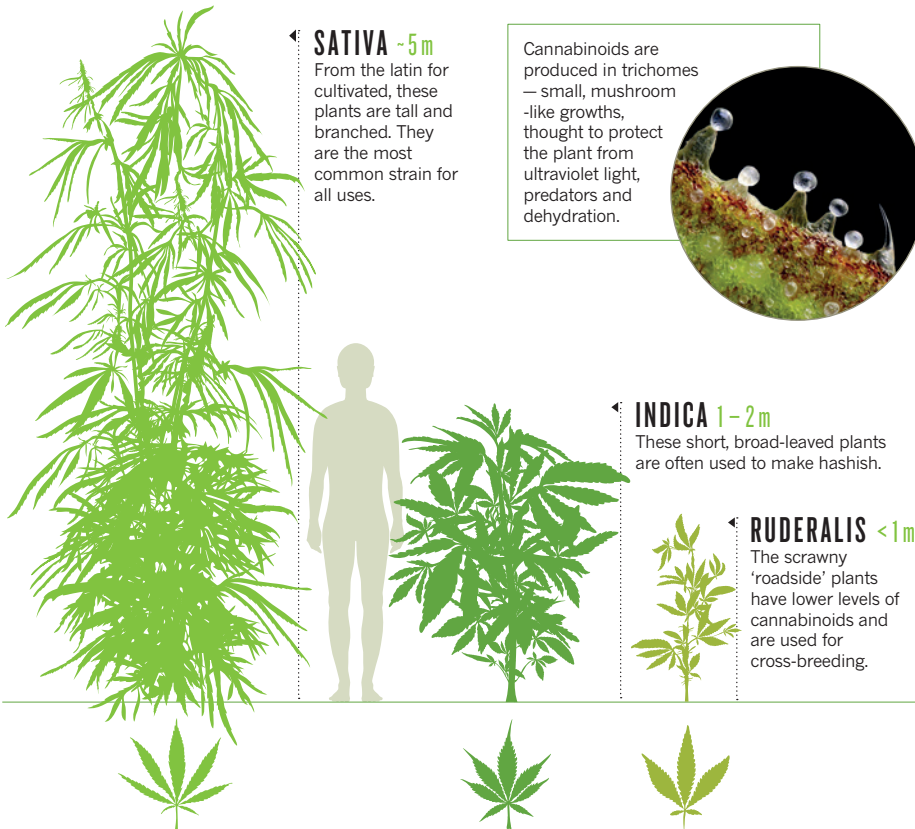
Material
The outer layer of the plant consists of long bast fibres, which can be used to make fabrics.



Medicine and intoxicants
Trichomes on leaves and buds (pictured) produce the plant's medically useful substances.

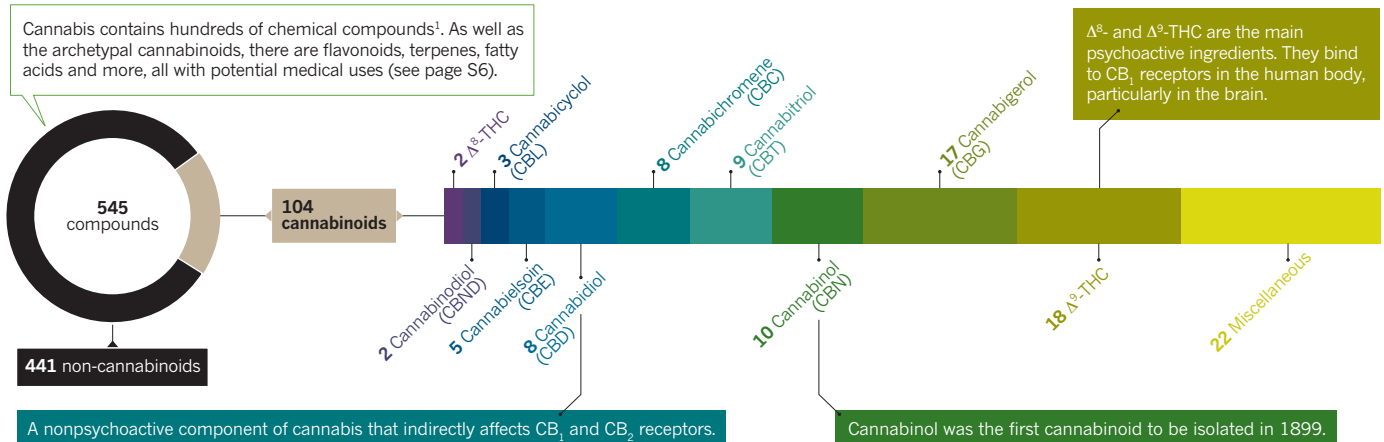


Food and cosmetics
Hemp seeds are technically nuts and contain more than 30% oil and 25% protein.



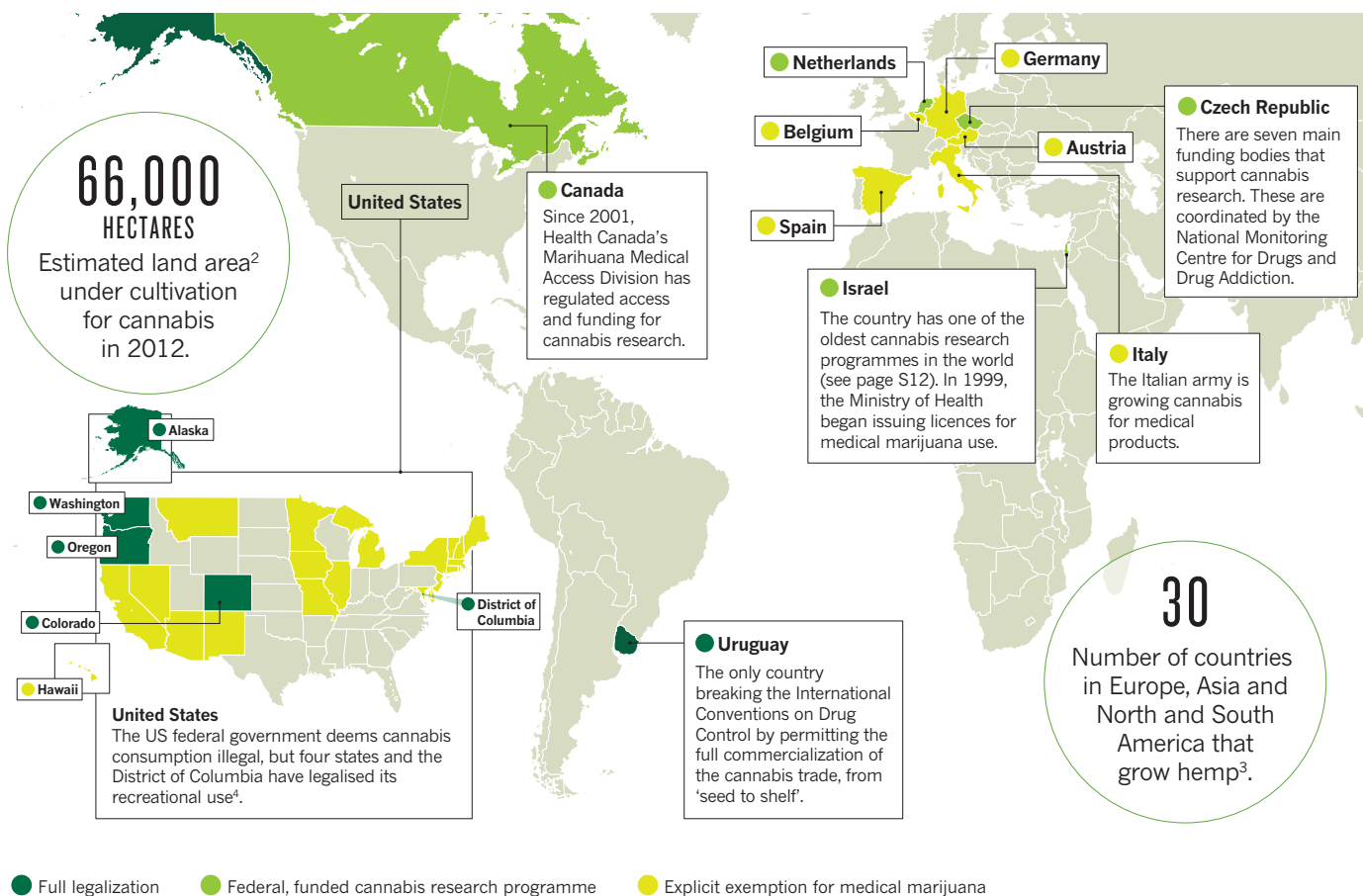
CHEMICAL CONSTITUENTS

Tetrahydrocannabinol (THC) is responsible for the mental high that can result from using cannabis. But there are many other cannabinoids and chemicals found in the plant, the roles of which are as yet unknown.



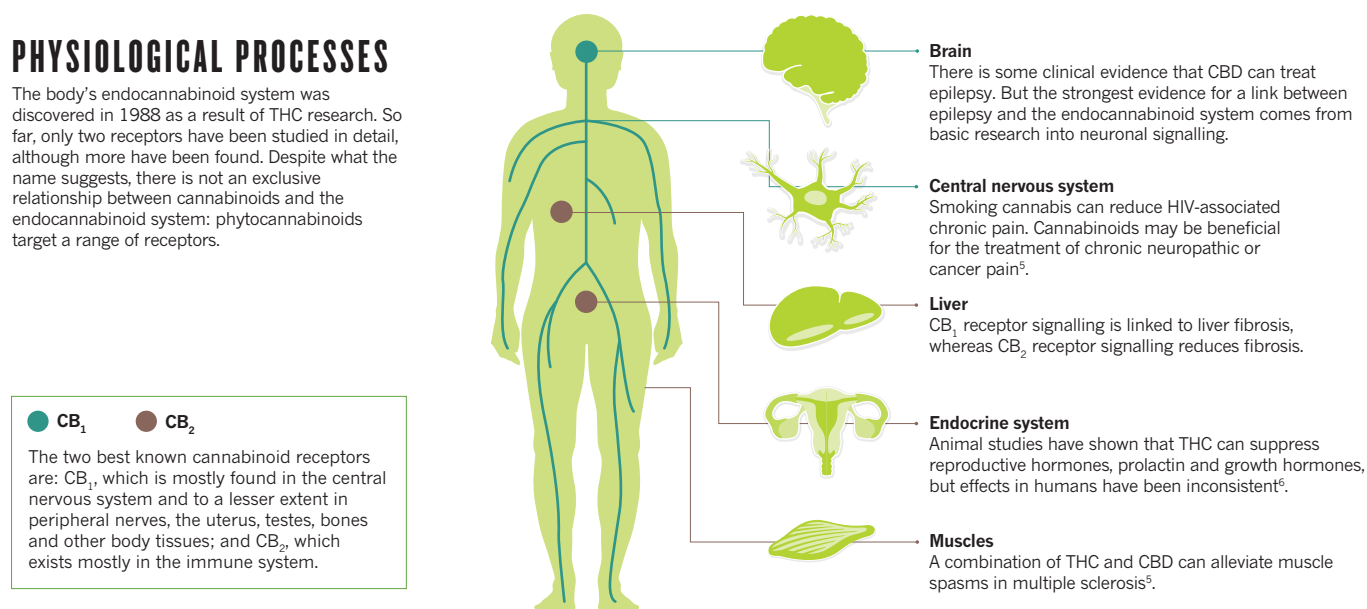
THE RESEARCH LANDSCAPE

The legal status of cannabis worldwide is in flux. One country and several US states have made herbal cannabis fully legal. Four countries have formal federal research programmes. Elsewhere, many countries have special exemptions for prescribed medical cannabis; others have decriminalized possession (not shown). Outside Europe and North America, however, severe punishments for even minor offences are common.



PHYSIOLOGICAL PROCESSES

The body's endocannabinoid system was discovered in 1988 as a result of THC research. So far, only two receptors have been studied in detail, although more have been found. Despite what the name suggests, there is not an exclusive relationship between cannabinoids and the endocannabinoid system: phytocannabinoids target a range of receptors.



1. Pertwee, R. G. (ed) *The Handbook of Cannabis* (Oxford Univ. Press, 2014). 2. United Nations Office on Drugs and Crime. World Drug Report 2014 (UN, 2014); 3. Johnson, R. Hemp as an Agricultural Commodity (Congressional Research Service, 2015). 4. *The Economist*. 5. Whiting, P. F. et al. *J. Am. Med. Assoc.* **313**, 2456–2473 (2015). 6. Brown, T. T. et al. *J. Clin. Pharmacol.* **42**, 90S–96S (2002).



CLAUDIA SCHAEFER PHOTOGRAPHY

Researchers at the University of British Columbia hope that by analysing cannabis diversity, they can determine whether it is one species or several.

BOTANY

The cultivation of weed

Researchers are getting closer to answering the centuries-old question of how to label cannabis varieties — a necessary step to bring the plant into mainstream agriculture.

BY LUCAS LAURSEN

Packets of cannabis seeds line the shelves of legal grow shops in Madrid. Many carry labels reporting the percentage of sativa and indica, two types of cannabis. Breeders often label plants that produce a more exciting high as sativa and plants that provide a more mellow feeling as indica, suggesting that cross-breeding tailors that buzz. The conceit is widespread. Botanist Jonathan Page at the University of British Columbia in Vancouver, Canada, says he sees the same at local grow shops.

For reasons that go beyond assessing the quality of the user experience, botanists such as Page are investigating the evolution and present-day diversity of cannabis. To do this, they must confront centuries-old taxonomic questions, including whether cannabis is one species, *Cannabis sativa*, with several subspecies or varieties, or if it is several distinct species, such as *C. sativa*, *Cannabis indica* and *Cannabis ruderalis*. "It's complicated taxonomically because of its intimate relationship with humans for long periods of time," Page says. People have long bred cannabis as a source of fibre, food and oil — as well as for its mind-altering effects

(see page S10). As governments relax cannabis laws, commercial growers want more clarity about the chemical properties and capabilities of the herb's many varieties. In parallel, regulatory bodies trying to establish a legal framework want to be able to classify whether a given type of plant is for fibre (hemp) or recreational or medical use (marijuana).

Demand for such information is pressing. Last year, the United States granted permission for farmers to grow hemp for research purposes. Several states, including Colorado, have legalized the possession and use of small amounts of marijuana, and are beginning to integrate the plant into the legal economy. Elsewhere, Uruguay has legalized cannabis and other governments are relaxing restrictions on its possession and use. As academic and commercial interest grows, governments and the research community will encounter a rising demand for taxonomic information to help resolve disputes, establish registered cultivars, and create reliable centralized databases of cannabis information. Botanist Ernest Small of the government agency Agriculture and Agri-Food Canada, says that talking about cannabis taxonomy "is

really talking about the ability of countries to rationally regulate important drugs and products."

BLURRED LINEAGES

Cannabis diverged around 27.8 million years ago from *Humulus*, the hop plant used to give beer its bitter and floral flavours, according to genetic analysis presented at the International Cannabis Research Society's 2010 meeting by botanist John McPartland and Geoffrey Guy of London-based GW Pharmaceuticals. Human influence on its diversity is more recent, but still stretches back millennia. The earliest archaeological evidence for human use of the plant comes from hemp ropes found in 10,000-year-old tombs in Taiwan. Cannabis now grows throughout much of the world, and humans have almost certainly had a role in shaping its many forms.

The plant is promiscuous, which confuses the species issue. Most known lineages seem to be capable of producing viable offspring from crosses with each other. Lines domesticated by humans may also have mixed with wild plants, blurring the taxonomic boundaries further.

The first modern taxonomist, the Swede Carl Linnaeus, used geographic origin and

SOURCE: GENBANK

sex organs to identify five variants of a single species, *C. sativa*. Later, the French naturalist Jean-Baptiste Lamarck used morphology and chemistry to distinguish *C. sativa* from a shorter, less fibrous and more psychoactive species, *C. indica*.

Debate continued throughout the twentieth century. The US botanist Richard Evans Schultes favoured a third species — *C. afghanica*. Small, however, disputed this, maintaining that the genus *Cannabis* had only one species, with several variants that had been selected for by humans. Small's expertise even took him into the courtroom to dispute lawyers' claims that the plant their clients had been caught with was a different — and hence unregulated — species from the *C. sativa* banned by law. Although many botany guidebooks and researchers now agree with Small's view, there is still debate — stoked whenever another scientist revisits or champions the arguments for multiple species — within the grower and user communities. "The issue is exaggerated and tends to mislead people," says Small. "I almost feel that it's better not to talk about it anymore." Yet, he and other cannabis researchers continue to encounter public demand for clarity.

Small continues to do research and has even provided authorities with a means of distinguishing between drug-type and non-drug-type cannabis — a chemical threshold. Thanks to Small's work examining the natural range of tetrahydrocannabinol (THC) concentrations, some governments are able to sidestep the taxonomy question by counting plants with less than 0.3% THC as hemp and those with more as marijuana. This has allowed the Canadian agriculture industry to cultivate groups of plants with stable characteristics and register these as formal cultivars of hemp without fear of running afoul of drug laws.

MOLECULAR AND GENETIC TECHNIQUES

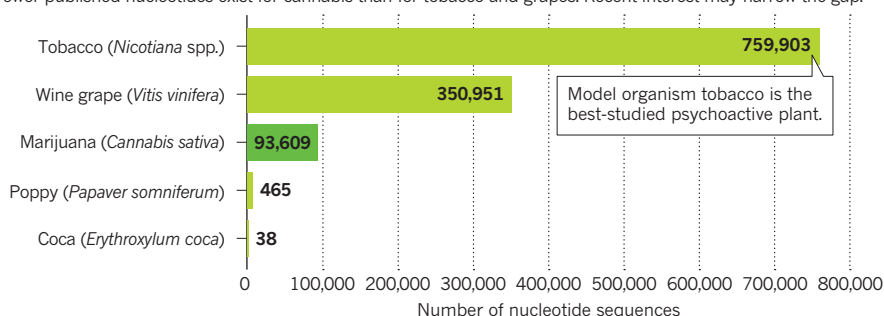
A chemical threshold is useful, but an official taxonomy would provide a clear and common language for researchers, regulators, growers and users to share information about the plants. In the past 20 years, researchers have turned to a variety of molecular and genetic techniques to tackle some of the questions that previous generations sought to resolve through morphology.

In 2004, biologists Paul Mahlberg and Karl Hillig, both then at Indiana University in Bloomington, analysed the enzyme-encoding genotypes of 157 sample varieties. Based on proportions of CBD and THC levels, they suggested that there are two species, *C. sativa* and *C. indica*, that contain six subspecies¹. Hillig later published a broader study identifying three species — adding *C. ruderalis* to the mix².

Genetic analysis, however, may not offer an immediate resolution to taxonomic debates: a 2003 study³ examining THC/CBD ratios identified five different lineages of cannabis, but all within one species. And in 2013, in perhaps the most comprehensive book on the subject,

GENOTYPING OUR HIGHS

Fewer published nucleotides exist for cannabis than for tobacco and grapes. Recent interest may narrow the gap.



botanist Mark Merlin of the University of Hawaii at Manoa and cannabis researcher Robert Clarke of the International Hemp Association in Amsterdam argued for three species of cannabis (*C. sativa*, *C. indica* and *C. ruderalis*), divided into a total of seven subspecies⁴.

Researchers continue to bring ever more sophisticated genetic tools to bear. In 2011, Page and his colleagues published a draft set of the DNA and RNA of a marijuana plant (*C. sativa*) and compared the RNA with that of a hemp cultivar. They found tantalizing differences in the expression of cannabinoid-controlling genes⁵. Botanist Nolan Kane of the University of Colorado Boulder is working with colleagues on a 'genetic map' that will involve complete DNA sequencing of some plants. They are also using a faster, cheaper method called genotyping by sequencing to study about 500 plants. By the end of 2015, they aim to have placed around 60,000 genes — about double the number reported by Page's group — onto the plant's 10 pairs of chromosomes. In addition, Kane's team has been working on determining the complete DNA sequence of 66 individual plants, with plans to extend to several hundred more. This work could be used to provide information about breeding new plants, and could also afford an "unprecedented" insight into the relationships between many of the major lineages of *Cannabis*, says Kane.

INDICA BY ANY OTHER NAME

Names and well-defined lineages matter because they help researchers to know what they are working with. "Many taxonomic studies and genetic studies work with *Cannabis* hybrids, and generate inconclusive results," McPartland says. Establishing groups of plants with stable features, each with some known characteristics such as certain THC and CBD levels or ideal growing conditions, could help pharmaceutical firms and others to exploit the plant (see page S6). What is more, without a clear taxonomy, existing lines with unique and useful traits may be neglected and even go extinct, he warns. McPartland's own research

suggests that some northern European strains have already disappeared.

This type of scientific omission might seem odd for such an apparently valuable plant. *Vitis vinifera*, a grape species used for making wine, has been subject to several genome-wide studies so far, and its cultivars are a matter of economic interest and national pride. And, tobacco (*Nicotiana tabacum*), although heavily regulated, is a model organism in basic biological research and has a well-documented pedigree (see 'Genotyping our highs'). But given cannabis's regulatory history and its stigma in many cultures, perhaps it is not surprising that there has been some reticence about its study. "This is a sensitive subject," Small says.

However, with an ever-growing number of jurisdictions permitting research and creeping towards cannabis commercialization, the need for a solid taxonomy is clear. Grow shops, with their labelled wares, are providing researchers with a bounty of specimens against which to test such 'folk taxonomies'. This year a study⁶ of 81 commercial marijuana samples demonstrated that the advertised percentages of sativa and indica show little correlation with the genetic reality. Unlike hemp, with its genetically stable registered cultivars, "in the marijuana world we don't have varieties or registered cultivars — we have things called strains", says Page. Strains are informally named by breeders and are not associated with a genotype in the same way that formal varieties or cultivars are. "You need to put a name to something to [research] it accurately," Page says.

"What is a species is a somewhat subjective concept," says Small. Whether a group of plants is a cultivar, a subspecies or a species may matter less than that everyone agrees on their evolutionary relationships. ■

Lucas Laursen is a freelance science writer based in Madrid.

- Hillig, K. W. & Mahlberg, P. G. *Am J. Botany* **91**, 966–975 (2004).
- Hillig, K. W. *Gen. Res. Crop. Evol.* **52**, 161–180 (2005).
- de Meijer, E. P. M. *et al. Genetics* **163**, 335–346 (2003).
- Clarke, R. C. & Merlin, M. D. *Cannabis: Evolution and Ethnobotany* (Univ. California Press, 2013).
- van Bakel, H. *et al. Genome Biol.* **12**, R102 (2011).
- Sawler, J. *et al. PLoS One* **10**, e0133292 (2015).



DRUG DEVELOPMENT

The treasure chest

Pharmaceutical research into the chemicals found in cannabis has so far supplied only one licensed medicine. But scientists think there could be hundreds more.

BY BRIAN OWENS

The annual meeting of the International Cannabinoid Research Society (ICRS) is a highly unusual scientific conference. It has been closed to all media since its inception 25 years ago, lending an air of mystery to the gathering of researchers who study the unique chemicals found in cannabis.

In a relaxation of the organization's long-standing policy, ICRS permitted *Nature* reporters to attend this year's conference, which was hosted by Acadia University in the tiny Canadian town of Wolfville, Nova Scotia. The tight-knit group of researchers are bound together by onerous government restrictions on their subject, and by their sufferance of lingering suspicions from other scientists that they are a bunch of hippies trying to get an illicit drug legalized.

"The status of cannabis as an illegal substance makes it difficult for some people to take it

seriously," concedes Mark Ware, a pain specialist at McGill University in Montreal, Canada, who focuses on the analgesic properties of cannabis.

But cannabis researchers are working hard to shed that image. On the whole they are not interested in the effects of smoking the plant. Their domain is the pharmacological study of the hundreds of chemical compounds in cannabis to determine how they could be developed into licensed pharmaceutical drugs to treat dozens of different conditions — while avoiding or minimizing the psychoactive effects. And they are slowly beginning to move these compounds from the laboratory to the clinic. "Stripped right down to the pure pharmacology, it's easy to make the case for cannabis as a medicine," says Ware.

TWO GREEN PATHS

Phytocannabinoids — a collection of more than 100 related chemical compounds found in cannabis — are the subject of most interest.

The primary, but not the only, targets for these compounds are the body's endocannabinoid receptors CB₁ and CB₂ (see 'A personable system'). Hence researchers have two avenues by which they can exploit the medicinal effects of cannabinoids. One strategy is to target the endocannabinoid receptors directly, by designing drugs that will activate or suppress them. The other is to harness the effects of the phytocannabinoids and turn these compounds into drugs.

The first approach has already resulted in one high-profile failure. In 2006, the European Medicines Agency approved a small-molecule anti-obesity drug called rimonabant (Acomplia) that suppressed appetite by blocking the CB₁ receptor. But, just two years later, it was withdrawn over safety concerns: people taking rimonabant had double the risk of developing

➔ **NATURE.COM**

For more on the endocannabinoid system see: go.nature.com/tudro5

psychiatric disorders, including depression and suicidal thoughts.

Despite this high-profile flop, Ware remains optimistic about targeting the cannabinoid receptors. “It continues to be a valid approach,” he says. “But it will be a number of years before we’ll see any new drugs emerge.”

The second approach — deriving drugs from the plant’s phytocannabinoids — offers a plethora of possibilities. Although researchers have identified more than 100 compounds, for decades pharmaceutical research started and ended with cannabis’s main psychoactive chemical tetrahydrocannabinol (THC; see page S2). There were a few studies into cannabidiol (CBD) as a treatment for epilepsy in the 1980s — led by chemist Raphael Mechoulam at The Hebrew University of Jerusalem who first discovered THC in the 1960s¹ (see page S12). But with the focus on THC and the widespread illegality of cannabis, this research was overlooked. It is only in the past few years that interest in CBD, as well as in other cannabinoids such as cannabigerol, cannabichromene and a THC variant called tetrahydrocannabivarin (THCV), has bloomed. “We’ve just scratched the surface,” says Jahan Marcu, director of research and development at the California-based company Green Standard Diagnostics, which helps labs purify and test cannabinoids. “There is lots of potential to explore other compounds that have great therapeutic indications.”

REVERSE DRUG DISCOVERY

Much of the research on the therapeutic benefits of cannabinoids started with anecdotal reports from people smoking cannabis to self-medicate for a range of ailments. Patients’ experiences were captured and used to inform pharmaceutical drug development. This method, which Ware calls “reverse drug discovery”, harks back to the way in which some of the most important drugs were found. The seventeenth-century observation that indigenous people in South America used the bark of the cinchona plant to treat malaria, for example, led to the discovery of quinine and the development of the first antimalarial drugs.

The approach has already led to the creation of synthetic phytocannabinoids. Dronabinol, for example, is a synthetic version of THC that is used to treat appetite loss in people with AIDS, and to relieve nausea associated with chemotherapy. Similarly, nabilone is used for nausea in patients undergoing chemotherapy.

So far, however, there is only one medication based on natural phytocannabinoids: a mouth spray called Sativex (nabiximols), made by GW Pharmaceuticals, based in

A PERSONABLE SYSTEM

Endocannabinoids are everywhere

The body’s endocannabinoid system is the pathway by which tetrahydrocannabinol (THC) exerts its psychoactive effects, and is the target for many of the plant’s other cannabinoids. It is intrinsic to a number of different processes, including appetite, memory, alertness, pain, inflammation and bone health, and stimulation of the endocannabinoid system is associated with the protection of healthy cells. “The endocannabinoid system helps us eat, sleep, relax, forget and protect our neurons,” says Jahan Marcu, director of research and development at California-based Green Standard Diagnostics.

Endocannabinoid receptors are spread throughout the body, and are believed to be more numerous than those of any other receptor system. Indeed, the fact that the endocannabinoid system is so widespread, and plays a part in so many different brain functions, could explain why the compounds found in cannabis seem to have no end of medical uses. Researchers have identified two main receptors so far: CB₁ and CB₂. CB₁ is found predominantly in the nervous system, but also in connective tissues, gonads, glands and organs; CB₂ is mainly found in the immune system.

The endocannabinoid receptors did not evolve just so that people could enjoy the effects of cannabis. The body produces its own cannabinoids: endocannabinoids, which are neuromodulators, meaning that instead of affecting just one neuron across a synapse, they diffuse throughout the nervous system and affect multiple

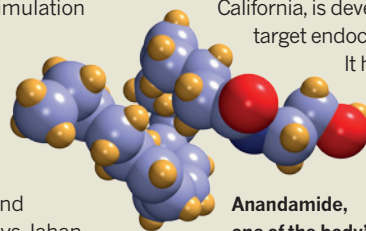
neurons (dopamine is another example of a neuromodulator). The two best understood endocannabinoids are anandamide, found primarily in the brain, and 2-arachidonoyl glycerol (2-AG), found mainly in the rest of the body.

Abide Therapeutics, based in San Diego, California, is developing treatments that target endocannabinoid levels directly.

It has developed a drug that increases natural levels of 2-AG in the brain by inhibiting monoacylglycerol lipase, a protein that breaks it down. Abide began enrolling participants in a phase I safety study in Belgium in July 2015.

The drug could potentially be used to treat neuropathic pain, neuroinflammation and even neurodegenerative diseases such as Alzheimer’s. “Modulating endocannabinoids in neurological diseases has real breakthrough potential,” says Abide president Alan Ezekowitz.

Raphael Mechoulam, a chemist at The Hebrew University of Jerusalem and the researcher who first identified THC, suspects that the endocannabinoid system is more than just a set of receptors. With crucial roles in processes such as memory, emotional response, learning and so on — he speculates that endocannabinoids are key to shaping people’s personalities⁶. “The body makes more than 150 endocannabinoid-like compounds — but why?” he says. “Is this one of the reasons individuals are different? Perhaps the different ratios of these compounds is part of what causes differences in personalities.” **B.O.**



Anandamide, one of the body’s cannabinoids

London, which is approved in 27 countries to treat spasticity associated with multiple sclerosis. Nabiximols is a whole-plant extract refined to contain about equal levels of THC and CBD. “Both THC and CBD have different pharmacologies that are complementary, in efficacy and safety,” says Stephen Wright, chief medical officer at GW Pharmaceuticals.

The company’s drugs are made from plants bred to have specific concentrations of the desired phytocannabinoids, says Wright. Cuttings are then taken from a mother plant to ensure each generation retains the same characteristics. “Once we have bred the plants that meet our needs,” he says, “we can control the chemical phenotype by fixing the genotype.” Such consistency is essential if a

company is to derive a pharmaceutical product that meets tough regulatory standards.

In addition to seeking approval for further uses of nabiximols — the drug is currently in phase III trials for cancer pain — GW is also exploring other cannabinoid agents. A CBD-based drug called Epidiolex (cannabidiol) is in phase III trials for two rare forms of epilepsy. And several other drugs are in phase II trials: THC and CBD to treat glioma brain cancer, THCV for type 2 diabetes and CBD for schizophrenia.

The results reported at the ICRS meeting reveal a vast array of further opportunities that are much earlier in the drug-development pipeline. Researchers presented data on the potential of phytocannabinoids to treat



At GW Pharmaceuticals' growing facility, cannabis clones are grown to ensure chemical consistency.

conditions such as acute and chronic pain, kidney disorders, Alzheimer's disease, opioid and nicotine dependence and post-traumatic stress disorder. "Cannabis is a bit like a treasure chest of compounds," says Roger Pertwee, a pharmacologist at the University of Aberdeen, UK, who has been studying cannabinoids since the late 1960s.

And it is not just cannabinoids in the chest. Cannabis also contains another class of compounds known as terpenes, which give the plant its characteristic smell (they are also found in cannabis's close relative,

hops). Some terpenes have been found to have anti-inflammatory, antibacterial, anti-anxiety or analgesic properties, says Ware. Various combinations of cannabinoids and terpenes could provide diverse therapeutic results, perhaps accounting for why people claim to experience different symptomatic relief from smoking certain strains (see page S4). "Whether [those effects] are therapeutically meaningful is unknown," Ware says.

Such combined interaction has been reported in the endocannabinoid system, as the 'entourage effect'. The body releases other chemicals at the same time as endocannabinoids, creating a stronger effect than that achieved with endocannabinoids alone. Many researchers think there is an equivalent entourage effect for phytocannabinoids. Pertwee gives THC and CBD as an example. THC acts through the CB₁ receptor as a painkiller, calming overexcited neurons, whereas CBD provides anti-inflammatory effects

through other routes. He also suggests that terpenes might modify the effects of phytocannabinoids. Ethan Russo, medical director of Phytects, Los Angeles, California, is studying this phenomenon. Russo's team is looking into whether combinations of complementary cannabinoids and other compounds, including terpenes, might work better than a single purified chemical².

There is still a long way to go before most of the prospective cannabis-based drugs make it into the medicine cabinet. The work presented at the ICRS meeting was almost entirely in animal models, with only a handful of clinical trials underway. "There have been lots of preclinical discoveries," says Pertwee. "Now we have to see whether they are hype or genuinely good ideas."

PRECISION MEDICINE

One of the big concerns about developing drugs from phytocannabinoids is whether it is possible to isolate the desired effects. THC, for example, is a potent painkiller, but comes with unwanted psychoactive side effects such as difficulty with recall. Peter McCormick, a pharmacologist at the University of East Anglia in Norwich, UK, may have found a way around that. He and his colleagues found that in mice that lacked a particular serotonin receptor THC still had painkilling effects, but did not impair memory.

THC targets the CB₁ receptor. And it turns out that the receptor missing in McCormick's mice, 5-HT_{2A}, also interacts with CB₁. When the team used a small peptide molecule to impede the 5-HT_{2A} and CB₁ interaction in normal mice, the same effect was seen³. He suggests that this could be a way to target the effects of cannabinoids more precisely. "It seems we can separate out the

'good' from the 'bad' effects of THC."

But even non-psychoactive phytocannabinoids that target CB₁ or CB₂ receptors can have unwanted side effects. "The endocannabinoid system is so ubiquitous, you can't target just one physiological process," says Ware. "It's hard to come up with clean drugs."

So researchers are looking for a way to modulate the effect of drugs that target this system. Endocannabinoid receptors have two binding sites: the main pocket and a secondary one — known as an allosteric site. A molecule that binds to an allosteric site changes the shape of the main receptor — altering the intensity of the effect of any molecule bound there⁴. Allosteric compounds, whether found in the body, in cannabis or elsewhere, could be used to fine-tune the effects of phytocannabinoid drugs, reducing unwanted side effects. Pertwee says that researchers have managed to identify a few allosteric endocannabinoids, but it is not known whether cannabis itself contains any. "We're now screening hundreds of compounds looking for good allosteric modulators," says Pertwee.

The other challenges facing cannabinoid research, however, are less tractable. Government restrictions on cannabis keep the field small, and make it difficult to get high-quality research materials, says Marcu. GW, for example, is the only reliable source of cannabigerol (which has shown promise as an antidepressant⁵), creating a bottleneck for research, he says.

And, despite the fact that researchers are working with purified phytocannabinoids, most of which have no psychoactive effects, they are still subject to the same restrictive regulations. This makes sending even small amounts of phytocannabinoids between labs in the United States all but impossible, says Marcu. And it means only a few pharmaceutical companies are willing to take the work forward into clinical trials.

Research is becoming easier as cannabis-based drugs, many of which contain no THC, become more accepted in mainstream medicine. But progress is still not fast enough for Mechoulam. "We knew 30 years ago that CBD lowered seizures for epilepsy, and it's the only thing that helps some kids," he says. Now 84, he has little patience left for the legal obstacles. "It's ridiculous," he says. "We're talking about the health of children here." ■

Brian Owens is a freelance science writer based in St. Stephen, New Brunswick.

1. Gaoni, Y. & Mechoulam, R. *J. Am. Chem. Soc.* **86**, 1646–1647 (1964).
2. Russo, E. *Br. J. Pharmacol.* **163**, 1344–1364 (2011).
3. Viñals, X. et al. *PLoS Biol.* **13**, e1002194 (2015).
4. Price, M. R. et al. *Mol. Pharmacol.* **68**, 1484–1495 (2005).
5. Musty, R. E. & Deyo, R. A. *Proc. Symp. Int. Cannabinoid Res. Soc.* 32 (ICRS, 2006).
6. Mechoulam, R. & Parker, L. A. *Annu. Rev. Psychol.* **64**, 21–47 (2013).

PERSPECTIVE



Close the knowledge gap

Nations with cannabis programmes should respond to a lack of research. Canada can be a leader, say **Jonathan Page** and **Mark Ware**.

When it comes to medical cannabis, Canada is both a leader and a laggard. Policy-wise, Canada is ahead of many other countries, having had federal regulations that allow patients to access herbal cannabis (dried leaves and flowers) with a doctor's authorization since 2001. Based on this early entry into medical cannabis, one would expect Canada to be at the forefront of research. Alas, this is not the case. As the number of patients accessing cannabis-based therapies has increased, research has not expanded. The opportunity to inform medical cannabis policy is slipping away.

The Canadian medical cannabis system continues to grow and evolve. The government, through Health Canada, has created a system to license producers to grow and distribute quality-controlled cannabis. Under this system, the patient population reached almost 24,000 in mid-2015, and around 4,000 doctors have prescribed cannabis. One would think that long-standing federal regulations and a large number of patients would mean that cannabis research is underway at many institutions in Canada. However, in the 14 years since the implementation of the first patient access programme, there have been only two federally funded clinical studies — a 2010 report that examined the use of smoked herbal cannabis to treat neuropathic pain¹ and a multicentre cohort study exploring one-year safety data². These studies were funded by Health Canada's Medical Marihuana Research Program, which was scrapped in 2006 as part of federal budget cuts. To our knowledge, no university laboratory in Canada has been licensed to grow cannabis for research purposes. The regulations that give patients access make no specific allowances for research.

Patients, doctors and producers have all expressed frustration with the regulations. Groups such as the Canadian Medical Association have made repeated calls for clinical trials and evidenced-based treatment guidelines. Physicians bemoan the lack of clinical data and the fact that herbal cannabis is not an approved drug; some also harbour suspicions that patients are seeking medical cannabis merely as a front for recreational use. Some cities are seeing a growth in the number of unlicensed dispensaries. And despite years of regulated access to the dried plant, access to cannabis extracts has only recently been mandated through a Supreme Court decision. Such uncertainties are not conducive to a well-functioning national medical cannabis programme that supports research and education.

Major gaps in knowledge persist. Fundamentally, the evidence base for the clinical use of herbal cannabis is thin. Although a recent systematic review³ found evidence for its use in treating chronic pain and spasticity, other claims were less well supported — creating a shaky foundation on which to base a treatment. Myriad other basic questions remain: for example, what are the pharmacological effects of diverse cannabis metabolites such as the non-psychoactive cannabidiol and volatile terpenoids? And we don't know whether individual strains of cannabis have different therapeutic properties⁴.

Why has there been so little progress? The Canadian government has not set medical cannabis as a public health priority, and so has provided insufficient research funding. At the same time, it has missed the opportunity to establish a national drug safety programme around medical cannabis. With a few exceptions, such as a newly initiated clinical trial focused on osteoarthritis, the private sector has yet to pick up the slack — partly due to a lack of patentable intellectual property. Research conducted by pharmacologists, chemists and plant biologists is an important complement to clinical investigations. However, pre-clinical cannabis research in Canada is impaired by delays and difficulties in obtaining research licences. To clear this logjam, funding should be channelled through a peer-reviewed cannabis research programme that can issue fast-track approvals for research licenses. These challenges are not unique to Canada. The same questions are being asked around the globe where national (such as Israel, the Netherlands and

Uruguay) and regional (various US states) medical cannabis-access regulations already exist or are being implemented. Ideally, these other governments will learn from Canada's experiences.

We call for a global initiative to identify and prioritize research needs around medical cannabis, alongside support to implement the research. At the United Nations level, we need a policy change to allow biomedical researchers access to cannabis and related materials, with the expectation that such liberalization will trickle down to national and regional programmes. To stimulate investment and enthusiasm for research, national medical cannabis offices need to be adequately resourced and given guidelines for streamlined and transparent review processes. Although public investment in cannabis research is important, harnessing funds from the burgeoning private

sector that is profiting from the sale of herbal cannabis could support many high-quality projects. The US state of Colorado provides an excellent model for just such an approach.

Promoting and facilitating research on cannabis is not an implicit acceptance of its medical value. Rather, it is a crucial response to an issue of global importance. The science of medical cannabis desperately needs to get out in front of the policy. It would be unforgivable if ten years from now we are still lamenting the lack of research despite widespread access to medical cannabis and a profit-hungry industry. ■

Jonathan Page is a plant biologist at the University of British Columbia in Vancouver, Canada. **Mark Ware** is a pain physician at the Alan Edwards Pain Management Unit at McGill University Health Centre in Montreal, Canada.

e-mails: jon.page@botany.ubc.ca; mark.ware@mcgill.ca

1. Ware, M. A. et al. *Can. Med. Assoc. J.* **182**, E694–E701 (2010).

2. Ware, M. A. et al. *J. Pain* (in the press).

3. Whiting, P. F. et al. *J. Am. Med. Assoc.* **313**, 2456–2473 (2015).

4. Russo, E. B. *Br. J. Pharmacol.* **163**, 1344–1364 (2011).

A potted history



EARLIEST EVIDENCE

~2700 BC

Cannabis sativa is thought to have been grown for at least 12,000 years, initially for fibre and grain. "The plant arose in Central Asia, but once people began growing it, it spread very quickly," says Ethan Russo, a psychopharmacologist at biotechnology firm Phytects, based in Los Angeles, California, and a historian of medical cannabis. The earliest use of cannabis as a medicine is attributed to the legendary Chinese Emperor Shen Nung (pictured), who is thought to have lived around 2700 BC. His teachings were passed down by word of mouth before appearing in writing in the *Shen Nung Pen-tsao Ching*, a second-century Chinese book of herbal remedies.

SCYTHIA'S STONERS ~450 BC

The Scythians of the Central Eurasian steppes erect a woollen tent, place a dish of red-hot stones inside and throw on hemp seeds, said Greek historian Herodotus. As the seeds begin to smoke, they inhale. "The Scythians enjoy it so much that they howl with pleasure," he recorded.



~700 BC

Archaeologists excavating the Yanghai Tombs in northwest China in the early 2000s identified one grave as that of a shaman buried 2,700 years ago. In the grave was a stash of well-preserved *C. sativa*. Later analysis of the plant remains confirmed the presence of the psychoactive tetrahydrocannabinol (THC). "This is the oldest physical evidence of pharmacologically active cannabis," says Russo.

POT BOILER ~190

The Chinese physician Hua T'o regularly anaesthetized his patients with a mixture of "hemp-boiling-compound" in wine before performing abdominal surgery.

WESTWARD BOUND 1545

Spanish colonists introduced cannabis to Chile, initially growing it for fibre. In 1611, English settlers took hemp to Jamestown, Virginia. Hemp went on to become an important crop in North America.

~390

The first evidence of medical use of cannabis came from a fourth-century burial in a cave near Beit Shemesh, 30 kilometres west of Jerusalem. Archaeologists excavating the cave in 1989 found the skeleton of a 14-year-old girl who had apparently died during childbirth. On her abdomen were burnt plant remains, which chemical analysis showed contained THC. The archaeologists concluded that cannabis had been burnt in a vessel and that the girl inhaled cannabis smoke during her efforts to deliver the baby.

SNACK ATTACK 1563

Portuguese physician Garcia da Orta, who lived in Goa, India, was the first Western observer to record the appetite-stimulating effects of cannabis: "Those of my servants who took it ... said that it made them so as not to feel work, to be very happy, and to have a craving for food."

EUROPEAN INTEREST

~800

In Western Europe, remedies were based on hemp, which has more non-psychoactive, but biologically useful, cannabidiol and less THC than Asian cannabis. Archaeological finds suggest that hemp was grown in Roman Britain for grain and fibre, but it was probably the later Saxons who used it as a medicine. The ninth-century medical text the *Old English Herbarium* advised pounded hemp for dressing wounds and a liquid concoction "for pain of the innards".

GETTING HOOKED 1689

The Royal Society's Robert Hooke was given, by a friend, "a drug from India called Bangué". Addressing the society, Hooke described it as "so like to hemp ... that it may be said to be only Indian hemp". He added that it might "possibly be of considerable use for Lunatics, or for other Distempers of the Head and Stomach".

Hemp



Modern medical interest in cannabis is traced to Irish physician William Brooke O'Shaughnessy. While in India, he saw how people used Indian hemp as a narcotic and medicine. Impressed, he tested it on animals before beginning trials in patients. O'Shaughnessy made extracts of cannabis resin and either rolled it into pills or dissolved it in alcohol to produce a tincture to treat conditions such as cholera, infantile convulsions and even tetanus. "O'Shaughnessy was of critical importance in introducing Indian hemp to British and North American physicians," says Russo.

1839

For thousands of years cannabis has been valued as a versatile herbal medicine. In the twentieth century, prescription gave way to proscription. Might this ancient remedy be about to regain its healing reputation? By Stephanie Pain

WAR ON DRUGS 1863

At the height of the American Civil War, a Union army soldier developed tetanus and gangrene after his shattered arm was amputated. He was treated with a tincture of cannabis and survived. Army doctors also prescribed cannabis with opium in an attempt to reduce the staggering death toll from diarrhoea and dysentery.

VIPER'S DRAG 1923

New Orleans was one of the first US cities to ban marijuana. Its popularity among the jazz musicians of Storyville, the city's historic red-light district, fuelled a moral crusade by organizations who saw the place, the drug and the music as a menace to society.



CALIFORNIA DREAMIN' 1996

California passed Proposition 215 (the Compassionate Use Act), which allowed the sale and medical use of cannabis for patients with HIV/AIDS, cancer and other serious and painful diseases.

HIGH HOPES 2011

The first draft genome of *C. sativa* was published (see go.nature.com/g1kffb), allowing researchers to explore its 100-plus cannabinoids and paving the way for the development of cannabis strains tailored to different medical uses.

1899

British chemists isolated cannabinalol, the first cannabinoid identified, but their discovery came just as medical cannabis was falling out of favour. Advances in chemistry made it possible to isolate and synthesize the active ingredients of medicinal plants, and tinctures gave way to drugs of guaranteed consistency. The hypodermic syringe accelerated the move to water-soluble drugs that could be injected for faster pain relief.

THE LAW STEPS IN

1925

An international treaty brokered by the League of Nations to control the opium trade was extended at the last minute to include cannabis. Signatories were required to control the trade in cannabis and prevent trafficking. The 1961 United Nations Single Convention on Narcotic Drugs clamped down still further, and a decade later the UN Convention on Psychotropic Substances made it all but impossible to carry out research on cannabis; only authorized people in supervised laboratories could work with it.

1964

In Israel, chemist Raphael Mechoulam isolated THC, kick-starting research into the plant's pharmacology (see page S12). "The main focus of research in the UK was in exploring the possible harmful effects of recreational cannabis," says neuropharmacologist Roger Pertwee of the University of Aberdeen, UK. "But gradually people got interested in the potential medical use of synthetic cannabis-like chemicals." Research led to nabilone and dronabinol, synthetic versions of THC, which were approved in the 1980s to suppress nausea during chemotherapy.

THE CANNABINOIDS WITHIN

1988

Researchers discovered a new receptor in the brain, named CB₁, through which THC exerts its psychoactive effects. "That led us to wonder if we have substances in our bodies that target this receptor," says Pertwee. The first of these so-called endocannabinoids, anandamide, was found in 1992. More followed, along with a second receptor (CB₂) in 1993. "The discovery that everyone has cannabinoids in their bodies led to a change in attitude," says Pertwee. "It made our research much more respectable."

1998

By the 1990s, growing numbers of people with conditions that failed to respond to prescription drugs were turning to cannabis. "We did a survey in the UK and the US, asking people with multiple sclerosis how they thought it helped them," says Pertwee. Based partly on the findings, an inquiry by the UK House of Lords concluded in 1998 that there was strong evidence that cannabis had a medical value, and in 2000 the government supported a trial of cannabis in multiple sclerosis.

2005

The first cannabis-based product Sativex (nabiximols) — a mouth spray of whole-cannabis extract, containing equal amounts of THC and cannabidiol — was given its first approval in Canada. The spray was developed by GW Pharmaceuticals, which was set up by Geoffrey Guy and Brian Whittle following the UK report. Today, nabiximols is approved in 27 countries to treat spasticity in patients with multiple sclerosis (see page S6).





BRENNAN LINSLEY/AP/PRESS ASSOCIATION IMAGES

Cannabis plants grown in Colorado were used to treat Charlotte's epilepsy, but researchers are heading to Israel to study the drug.

ISRAEL

Research without prejudice

How one Mediterranean country is pushing the frontiers of medical cannabis knowledge.

BY EMILY SOHN

Alan Shackelford is intent on finding out why some of his patients respond so well to cannabis. But despite living in Colorado, the US state with some of the most liberal medical marijuana laws, he has had to travel to Israel to continue his research.

Shackelford's road to the Mediterranean nation started in 2012. While working in occupational medicine and injury rehabilitation private practice, he got a call from a mother whose 5-year-old daughter Charlotte was having 300 seizures a week and not responding to treatment. The family were desperate for help. They had heard that medical marijuana was being used to treat epilepsy, but had been turned away by doctors when they asked for the treatment for Charlotte. Although Shackelford had finally agreed to treat his older patients with cannabis a few years earlier, he was particularly reluctant to give the herb to such a young child. But, after digging into the literature, Shackelford agreed to treat Charlotte with a specific strain high in cannabidiol (CBD), which a friend of the family converted into an oil extract.

Now 8, Charlotte is thriving. She takes the oil every day and has just one seizure every month or so, Shackelford reports. He has seen

other, similar stories, but such case reports and testimonials do not constitute peer-reviewed evidence. However, when he looked into getting permission for a trial, he was overwhelmed by the bureaucracy involved. At a federal level, cannabis is classified as a schedule 1 drug, meaning that it has no known medical value. Unless the study looks at the harm the drug might cause, permission for cannabis research can be harder to obtain than that for heroin or cocaine, says Shackelford. "There is a bias against doing trials here that might show a benefit."

Frustrated, he went to Israel — one of only a few countries with a national medical cannabis research programme. Shackelford was attracted by the country's 50-year history of study into potential uses for the drug, as well as a supportive regulatory atmosphere that is not found anywhere else. "The attitude towards research in Israel has always been different and not coloured by prejudice or propaganda," Shackelford says. Reputable researchers who want to study cannabis are not simply dismissed, he says, "which is often the case in other countries, including, notoriously, the US".

Other researchers and entrepreneurs are, like Shackelford, turning to Israel to

further research into cannabis. "It's a topic in which, maybe surprisingly, Israel is pushing ahead," says Raphael Mechoulam, a chemist at The Hebrew University of Jerusalem.

GOOD PARENTING

The modern era of cannabis research started in Israel, spearheaded by Mechoulam (see page S10) — often called the father of medical cannabis. In fact, Shackelford's decision to treat Charlotte was influenced by Mechoulam's research into CBD (J. M. Cunha *et al. Pharmacology* **21**, 175–185; 1980).

Today, Israel is one of many places that boasts a broad supportive atmosphere for cannabis: some 75% of the population back its medicinal use, and in 2013 the Orthodox rabbi Efraim Zalmanovich ruled that medical cannabis was kosher.

Israel also seems to nurture an entrepreneurial spirit, which is apparent in Mechoulam's story. Originally from Bulgaria, Mechoulam began investigating cannabis in the 1960s while working at the Weizmann Institute of Science in Rehovot. He was attracted by the mystery: although the active constituents of coca leaves and opium were known, cannabis was still largely unstudied.

Before he could start his research, Mechoulam needed to procure some cannabis.

➔ **NATURE.COM**

Read more about the research challenges at: go.nature.com/tg6d9v

Unsure how to get it, he asked his superior if he knew anyone in the police department who could give him some of their confiscated supplies. The administrator called the police, and Mechoulam recalls hearing someone on the other end of the line ask, “Is he reliable?” Once assured, the police invited him to come by. “I went there and picked up five kilos of hashish,” Mechoulam says. However, he soon found out that this was not the correct procedure. “It turned out we had broken the law, and the police had broken the law. We should have gotten all kinds of permits.” But he was not punished. “I went and apologized,” Mechoulam says. He filled in the proper paperwork: “And for years, I continued to get legal cannabis from the police.”

With a reliable supply of cannabis, Mechoulam and his team made rapid progress. In 1963, they determined the structure of CBD. The following year, they isolated tetrahydrocannabinol (THC), the main psychoactive substance in cannabis. Mechoulam also helped to discover anandamide, a naturally occurring chemical in the brain that binds to the same receptor as THC, sparking interest in the ‘endocannabinoid’ system (see page S6).

Currently, says Mechoulam, there are about ten research groups in the field in Israel, working on cannabis as a treatment for conditions such as post-traumatic stress disorder (PTSD), epilepsy, chronic pain, rheumatoid arthritis, fibromyalgia and Crohn’s disease (see page S15). At 84 years old, Mechoulam is still an active researcher and is currently planning a trial of cannabis in people with brain cancer. His adopted country has long celebrated him, and he is free of the reputation-threatening stigma that is often attached to cannabis researchers elsewhere.

This is in contrast to the saga of US psychiatrist Sue Sisley, who experienced the negative side of cannabis research. Sisley is currently waiting to begin a phase II trial of cannabis for veterans with PTSD. The US Food and Drug Administration (FDA) approved the study in April 2011, but getting the go-ahead to do clinical trials also required approval from the National Institute on Drug Abuse (NIDA), the Drug Enforcement Administration and the Public Health Service (this stage was dropped in June 2015 to reduce bureaucracy). NIDA and the Public Health Service rejected Sisley’s plan. Although she finally secured approval for a revised proposal in the spring of 2014, soon after that she lost her job at the University of Arizona in Tucson. She recalls that, during multiple meetings and phone calls, administrators told her that they feared the university would lose its federal funding if it allowed cannabis to be studied on campus;

“They were actually there to very sincerely look into the matter of when drugs are useful.”

the institution maintains that its decision not to renew her contract was unrelated to the politics of medical marijuana research.

Sisley flew to Israel earlier this year to investigate the potential of moving the research there, and describes the experience as eye opening. “It was a joy to be there,” she says. “None of the researchers perceived marijuana as an impediment to science. They just viewed it as another study drug.” For the time being, Sisley has decided to remain in the United States as an independent researcher. Her study has received approval from an independent institutional review board and, with her team, she has secured funding and found space in which to conduct the study. She is now waiting for four cannabis strains from NIDA to finally start her trial.

SMALL IS BEAUTIFUL

Working in a relatively small country with a limited number of research facilities increases the chance that scientists will talk across disciplines, particularly for researchers in a niche field such as cannabis. Immunologist Ruth Gallily recalls how her interest in the drug began after a chance encounter with pharmacologist Esther Shohami some 15 years ago while they were both working at The Hebrew University of Jerusalem. Shohami described her experiments in which mice with head trauma showed great improvement when treated with a CBD-derivative (D. Panikashvili *et al. Nature* **413**, 527–531; 2001). Gallily asked to see the animals’ brains and was surprised to see that those treated with the derivative showed suppression of a key inflammatory protein. This led her to study the anti-inflammatory effects of CBD and its potential to treat conditions such as diabetes, heart disease and multiple sclerosis. “In a small place, you meet people again and again,” Gallily says. “You begin to get ideas from others.”

Timna Naftali, a gastroenterologist at Meir Hospital in Kfar Saba, says that her experience of seeking approval for two trials of medical marijuana as a treatment for Crohn’s disease was smooth — a process that could have been much more difficult had she been working in the United States or even Canada. “I was expecting an approach like that of a drug authority that would say, ‘no’ to everything because they are there to stop people from using drugs,” says Naftali. “But they were actually there to very sincerely try and look into the matter of when drugs are useful and when they are not.”

Israeli technology companies are also involved in medical cannabis, developing delivery systems such as sustained-release pills and vaporizers. The country is even attracting foreign companies, who want to base their research and development (R&D) centres in Israel. Eyal Ballan is co-founder and chief scientist at drug-development company Canabics Pharmaceuticals, based in Bethesda, Maryland, the R&D arm of which is in Israel.



Alan Shackelford (left) and Raphael Mechoulam.

Ballan says that he has received calls from researchers in places such as Brazil, Germany and the United States. “Sometimes they want the product itself,” he says. “Sometimes they need help with regulations to promote medical cannabis in their states. Sometimes they wish to test their own products in Israel.” This experience is echoed by Boaz Wachtel, co-founder of Phytotech Medical — an Australian cannabis research company that also bases its R&D in Israel. Once every couple of weeks, he says, he hears from researchers and executives looking for advice or seeking research collaborators.

Israel’s approach to cannabis is more liberal than those of most countries, but it is far from a free-for-all. The drug remains illegal for recreational use (although there are signs that this may be changing). Israel also refuses to export cannabis to other countries, despite plenty of interest. Many researchers consider that this kind of balanced approach may be an important factor in why cannabis research in Israel is taken seriously; the herb is treated as a drug that needs to be studied in order to be safely used, just like any other. “We have to know exactly what amounts of CBD and THC people are getting,” says Mechoulam. “These things have to be well-regulated.”

These policies have served the country well. “The Israeli national medical cannabis programme is considered a success,” says Wachtel. “I see a great movement, from the United States, from Canada, from many other countries, of people who want to get into this field. We are awash with newcomers.” ■

Emily Sohn is a freelance journalist based in Minneapolis.

PERSPECTIVE



Be clear about the real risks

The assertion that cannabis use can cause schizophrenia is not borne out by the evidence, says **Matthew Hill**.

The 1936 film *Reefer Madness* depicted cannabis as a drug that provoked uncontrollable insanity, leading to manslaughter, suicide and attempted rape. This was a ridiculous characterization of the effects of cannabis, but there is a long history of associating the drug with psychotic disorders. In research terms, the first evidence came from a 1987 study, which found that Swedish conscripts had an increased risk of developing schizophrenia if they had consumed cannabis more than 50 times in their life¹. This finding has been replicated, implying, at the very least, an intricate relationship between cannabis use and schizophrenia². The nature of this relationship is still a matter of debate and is not as clear as some researchers or policymakers would suggest.

One interpretation is that cannabis is an instigating factor in the development of schizophrenia. Some researchers have argued that removing cannabis, particularly high potency strains, from society would reduce the prevalence of the disease³. Although this may seem alluring from a drug regulatory standpoint, from a scientific one we need look at the evidence. The history of cannabis use in the Western world stands as an enlightening social experiment. Before the 1960s, cannabis use in Europe and North America was relatively uncommon; today, use varies between countries, but in certain regions upwards of 20% of the adolescent population use the drug. If cannabis is causally related to the development of schizophrenia, then it would be expected that the incidence of the disease would have increased significantly with increased use of the drug. Yet schizophrenia rates since the 1960s have remained stable worldwide, and in fact even declined slightly in the West between the mid-1960s and the mid-1990s⁴. Although changes in diagnostic practices and disease classification may have contributed to this drop, if cannabis does induce schizophrenia, then we would expect to have seen a jump in the number of cases. There also seems to be no difference in schizophrenia rates between countries where cannabis use is prevalent and those where its use is rare. Again, if the drug were an instigating factor alone, we would expect to see differences in the population data.

Other evidence purporting to support the causal hypothesis is also inconclusive. Clinical studies have shown that pure tetrahydrocannabinol (the psychoactive constituent of cannabis) can produce an acute psychotic state. But these states are transient and do not lead to mental illness. It is also known that people with schizophrenia consume cannabis more than the general population⁵. Although cannabis may worsen schizophrenic symptoms such as delusions and hallucinations, it might also mitigate negative symptoms, such as anxiety and social withdrawal — explaining why people with schizophrenia would want to use it⁵. As with any correlation, there is the possibility that a third variable mediates the relationship. The finding⁶ that genetic variance could predispose a person to schizophrenia and also increase risk of cannabis use could explain the co-occurrence of these variables on a biological basis.

What does seem to be clear is that heavier than average cannabis use, particularly in early adolescence, can accelerate the onset of

schizophrenia⁷. Although this seems damning at first glance, it is less so when examined more closely. A reasonable interpretation is that individuals with schizophrenia, in their attempt to self-medicate, tend to use the drug more frequently than the general population. A vicious circle could develop whereby an adolescent in the early phases of psychosis begins to use cannabis to mitigate some aspects of their developing symptoms, but in fact speeds up disease onset. In this sense, cannabis would be one influence in an already developing illness as opposed to a stimulus that induces the development of the disease itself.

In support of the hypothesis that cannabis only triggers the onset of schizophrenia, gene variants have been identified⁸ that predict the development of schizophrenia in response to cannabis use. This suggests that cannabis promotes the development of schizophrenia only in people with a specific biological predisposition. Imagine that the disease is like a campfire: adding fuel to a pile of sticks has little effect,

but throwing fuel on a weakly burning fire will increase its strength. Regardless of whether fuel is added, the embers will continue to burn. This hypothesis would also explain the epidemiological data: that higher rates of cannabis use are associated with schizophrenia, but cannabis use does not affect disease rate at a population level. Importantly, this would mean that cannabis does not induce schizophrenia in non-vulnerable individuals. The distinction between the causal and the trigger hypothesis is significant in the message that is conveyed to the public. The former suggests that cannabis use alone can cause the disease, whereas the latter indicates that cannabis is merely a risk factor for someone who would probably develop schizophrenia anyway.

Clearly, understanding the nature of the risk of schizophrenia is important when developing social policies surrounding cannabis. Education about the drug's effects on mental health should highlight the association of cannabis use with schizophrenia. But scientists should be careful with the language that they use, particularly when presenting this relationship to the public. It is important to ensure we do not confuse correlation with causation and incite another *Reefer Madness*-style panic. By offering careful, evidence-based interpretations of the data, scientists can effectively contribute to policy decisions related to cannabis use and mental health. ■

Matthew Hill is a cannabinoid neuropharmacologist in The Hotchkiss Brain Institute, University of Calgary, Alberta, Canada
e-mail: mnhill@ucalgary.ca

1. Andreasson, S., Engström, A., Allebeck, P. & Rydberg, U. *Lancet* **330**, 1483–1486 (1987).
2. Hall, W. & Degenhardt, L. *World Psychiatry* **7**, 68–71 (2008).
3. Di Forti, M. et al. *Lancet Psych.* **2**, 233–238 (2015).
4. Warner, R. *Schizophr. Bull.* **21**, 483–500 (1995).
5. Kolliakou et al. *Int. J. Dev. Neurosci.* **29**, 335–346 (2011).
6. Power, R. A. et al. *Mol. Psych.* **19**, 1201–1204 (2014).
7. Large, M., Sharma, S., Compton, M. T., Slade, T. & Nielssen, O. *Arch. Gen. Psychiatry* **68**, 555–561 (2011).
8. Di Forti, M. et al. *Biol. Psychiatry* **72**, 811–816 (2012).



Proponents' claims about the medical benefits of marijuana are outpacing the evidence.

MEDICAL MARIJUANA

Showdown at the cannabis corral

Researchers are gathering clinical data for medical marijuana against a backdrop of deregulation and opportunism.

BY MICHAEL EISENSTEIN

Mary Jane Rathbun was beloved among patients with AIDS at San Francisco General Hospital in the late 1980s and early 1990s. As a volunteer, she took them for X-rays, filed their prescriptions — and supplied them with marijuana-laced brownies to alleviate their debilitating pain and wasting symptoms.

That service resulted in 'Brownie Mary' being arrested on multiple occasions. But patients got relief from her deliveries, and doctors both at her hospital and elsewhere were taking note. Barth Wilsey, then a pain research fellow at the University of California, San Francisco, also heard patients claiming they benefitted from this illicit substance. "I had a number of patients going to a dispensary in Oakland who told me that they were getting more relief from marijuana than from the medicine I gave them," he says. Inspired by Brownie Mary, Donald Abrams, an oncologist at the same hospital, tried mounting a clinical trial for medical marijuana as a treatment for patients with AIDS, but his efforts yielded

years of frustration. He recalls a 1997 conversation with Alan Leshner, then head of US drug research agency NIDA. "He told me that they're the National Institute 'on' Drug Abuse, not 'for' Drug Abuse," says Abrams.

Much has changed since then. Abrams, Wilsey and others have steadily accumulated data that suggest that medical marijuana has a clinical benefit for treating chronic pain, and the barriers that previously thwarted research are eroding. In June, the US government lifted a major hurdle in the grant-review process for cannabis research by removing the requirement for Public Health Service (PHS) review, and NIDA is beginning to offer researchers a broader range of cannabis strains. Meanwhile, several countries and nearly half of all US states authorize medicinal marijuana use. These changes create opportunities for researchers to embark on more robust trials and to directly observe whether the herb helps patients.

However, medical marijuana may also be a victim of its own success. Marijuana laws make assumptions about medical benefits that outpace the evidence, and in some cases create

markets for retailers to peddle unproven medicines — touting, for example, cannabis-derived oils to 'cure your own cancer'. Arno Hazekamp, who heads research and education at Bedrocan in Veendam, the Netherlands — the Dutch government's official provider of medical cannabis for more than a decade — dubs such opportunists "cannabis cowboys". "Everybody is selling stuff, but the real professionals who are supposed to do the job are not there yet," he says. As a result, some of cannabis's scientific supporters fear that poorly planned deregulation could undermine efforts to establish medical legitimacy for this controversial crop.

THE PROMISE OF POT

Much of the seminal clinical research into medical marijuana (that is, in its herbal form) was performed under the auspices of the Center for Medical Cannabis Research (CMCR), funded by the state of California between 2000 and 2003. With CMCR support, Abrams and colleagues conducted a randomized trial showing that 52% of cannabis users reported a meaningful reduction in HIV-associated neuropathic pain compared with 24% of the control group¹. With the help of a funding initiative from Health Canada, pain specialist Mark Ware and his team at McGill University in Montreal reported similar findings in patients with neuropathic pain as a result of a range of conditions².

Early data for several conditions are intriguing, but limited. After learning that her patients with Crohn's disease were finding relief with cannabis, gastroenterologist Timna Naftali at Israel's Meir Medical Center in Kfar Saba and her team investigated whether medical marijuana can induce remission of the disease. They found that it significantly improved appetite and pain symptoms, and allowed some patients to end their dependency on steroids. However, cannabis had no clear effect on the inflammatory processes that underlie Crohn's disease, and did not result in remission³. "The question is whether they were feeling better because it is a painkiller or reduces stress, or whether it really does something to inflammation," says Naftali.

Other findings are more tentative or anecdotal. Parents of children with severe epilepsy have reported dramatic improvements from cannabis oils derived from strains rich in cannabidiol⁴ (CBD) — one of the plant's non-psychoactive chemicals. Studies of marijuana for the treatment of mental health conditions such as post-traumatic stress disorder (PTSD) have yielded ambiguous results: some suggesting that individuals with PTSD may self-medicate to achieve restful sleep — whereas other data suggest that marijuana can actually disrupt healthy sleep patterns. "The PTSD literature is a mess," says Mitchell Earleywine, a psychologist at the University of Albany, State University of New York. He queries whether the different results arise from patients using different



Medical marijuana is sold in dispensaries, but some people fear that it will be used recreationally.

marijuana strains, but is also concerned that the herb may offer limited long-term value to these patients by acting as a “Band-Aid” for a condition that could benefit more from direct psychiatric intervention. “Personally, I’d rather see these folks do behavioural treatments.”

QUALITY CONTROL

The empirical basis for medical marijuana is thin. A systematic review of clinical studies published in June found “low-quality evidence” for the use of cannabis in almost all conditions, with the exception of chronic pain and muscle spasticity in multiple sclerosis⁵. The trials were often confounded by poor design or execution, or failed to objectively demonstrate clinical benefit. Most studies examined isolated cannabinoids, either natural or synthetic; data from trials in which the plant is smoked or vaporized were especially limited.

Even the best trials were narrow in duration and scope. “These studies do not extend beyond more than two to four weeks, and are limited in the range of doses that are available in terms of THC [tetrahydrocannabinol] potency,” says Ware. Finding a placebo for a trial of a drug with well-known psychoactive properties can be problematic. Margaret Haney, a neurobiologist at Columbia University in New York, notes that there is a considerable expectancy effect — preconceptions about cannabis skew how users respond. “It gets you high, and if you believe it’s going to cure everything under the sun, then when you smoke it you feel like it cures everything under the sun,” she says. And for conditions such as pain or appetite control that rely on user reports, expectations will have a big impact. Possible solutions include cannabis strains that are devoid of psychoactive THC, or placebos laced with mild sedatives.

The barriers facing even well-designed trials

can be exceedingly high. In the United States, cannabis is a schedule 1 controlled substance (it has no currently accepted medical use). Studying the drug requires researchers to grapple with an alphabet soup of agencies, including NIDA, the Food and Drug Administration (FDA), the Drug Enforcement Agency (DEA) and until recently the PHS. The cost of large-scale, long-term human studies can also be prohibitive. These bureaucratic and financial barriers have led to something of a stalemate, where researchers have no robust clinical data, but also lack the wherewithal to produce any. “If we just had a large funding source for even one good randomized controlled trial on a disorder that people care about, that would do the trick,” says Earleywine.

US researchers are now seeing some hope. The lifting of the PHS review requirement should accelerate grant review. “The first time I studied marijuana, that stage took 18 months, and I couldn’t go to the DEA or FDA until they concluded their review,” says Wilsey. And NIDA’s expansion of the range of strains available through its one authorized growing facility (at the University of Mississippi) means that researchers will have access to plants with THC levels that more closely resemble those found in street or dispensary cannabis as well as strains with higher levels of CBD.

However, the rapid policy changes now underway in North America will not necessarily help efforts to evaluate cannabis’s medicinal value. Canada first established a legal medical cannabis framework in 2001, but the initial system proved impractical. “Health Canada put this very complicated programme into place, and only a few hundred people went through the process and got approved,” says Benedikt Fischer, a public health researcher at the Centre for Addiction and Mental Health in Toronto. A series of court decisions led to

the launch of a reformed programme in April 2014; now the government oversees cannabis production but not prescription access, which is at the discretion of individual doctors. “It’s completely wide open — in many ways, arbitrary,” says Fischer.

If Canada’s system is arbitrary, the situation in the United States could be described as bewildering. At one extreme are states such as California, where patients can obtain almost any strain of cannabis by finding a doctor willing to provide a recommendation — merely a formality in some locales. By contrast, Minnesota offers restricted access to limited quantities of cannabis extract in liquid or pill form through state-run dispensaries for a narrow range of disorders. So far, 23 states and the District of Columbia have each come up with distinct approaches for providing medical cannabis — all in open defiance of federal law (see page S2). “There is no standardization,” says J. Michael Bostwick, a psychiatrist at the Mayo Clinic in Rochester, Minnesota. “I guess we could have ‘best practices’ for states defying federal law, but that seems like adding convoluted craziness to what is already a crazy situation.”

North of the border, the Canadian Medical Association and other professional organizations oppose the medical marijuana policy because they say it is based on insufficient evidence. Pharmacologist Harold Kalant of the University of Toronto notes that many doctors resent being appointed gatekeepers for this unproven drug — but even as a critic of the Canadian system, he believes the medical community has a responsibility to engage with this issue. “Who is better equipped to do this?” asks Kalant. “But this means they have to buckle down and do some reading and make themselves knowledgeable.” And that includes keeping abreast of the latest thinking about potential risks (see ‘Calculating the costs’).

The problems associated with implementing a medical marijuana law are in part attributable to the fact that this decision, which requires sophisticated scientific understanding, has been given to the public.

“If the majority of voters are in favour of overall legalization, then let’s just make it consistent across states,” says Haney, “but don’t have people vote on whether it’s a medicine.” Among jurisdictions that limit medical marijuana to specific conditions, there is little consensus on the best approach, whereas those that have adopted more open-ended policies have created a milieu of easily-obtained permits and abundant dispensaries that are largely indistinguishable from outright legalization. “People are saying that medical cannabis is just a scam and a cover for recreational use,” says Ware. “I worry that the message that there

“Those who are very ill can get cannabis anyway, and don’t need a study.”

may be medical value gets lost.”

This also gives free rein for cannabis cowboys to make outlandish claims. Abrams is no longer able to keep up with the e-mails and calls about ‘miracle cancer cures’. “I’m a little frustrated,” he says. “There’s just no data to support this.” Other companies offer products that piggyback on promising preliminary research — but with little clinical proof, regulatory oversight or quality control. Hazekamp cites a recent FDA investigation of extracts from CBD-rich cannabis strains, which are being touted as a potential treatment for severe epilepsy, cancer and diabetes. “Almost none of the products complied with their labelling,” he says, “and some contained no CBD at all.”

HIGH TIME FOR RESPONSIBLE RESEARCH

These legal experiments could yet bear scientific fruit. Colorado is using money raised from its recreational marijuana industry to support nine clinical research grants, several of which are for observational studies of real-world patient use. In Canada, Ware helped establish the Quebec Cannabis Registry earlier this year, which will track patient health and outcomes over four years. Fischer and colleagues are developing a similar programme to monitor patients in Ontario.

There is a risk, however, that these efforts could be confounded by recreational users gaming the system. Hazekamp notes that in the Netherlands, where recreational cannabis is widely available, the proportion of people accessing medical marijuana is less than one-tenth that of some US states⁶. Canada’s medical use is also booming, and Fischer likens the situation to Prohibition in the 1920s and early 1930s, when alcohol consumption in the United States was restricted to religious and medicinal uses. “Suddenly thousands of people needed medical alcohol — and it’s not like there was a sudden epidemic of ill people,” he says. Observational data are also no substitute for randomized controlled trials — and ironically, medical marijuana access may impede recruitment for the trials needed to prove its value. “Those who are very ill can get cannabis anyway, and don’t need a study,” says Naftali.

Veteran cannabis researchers are finding new opportunities to strengthen their case. Wilsey’s team is embarking on a neuropathic back-pain trial spanning 8 weeks, with more than 100 patients, while Abrams is working with NIDA’s new cannabis strains to assess the combined effects of THC and CBD in modulating pain in sickle-cell anaemia — a condition associated with especially debilitating and difficult-to-manage pain. Naftali is conducting a 50-patient Crohn’s disease trial that will closely examine physiological markers as well as patient-reported symptoms. “We’ll be doing endoscopy before and after patients take cannabis, to see whether there’s any real difference in the inflammation,” she says.

Funding opportunities are also attracting

CALCULATING THE COSTS

How risky is weed?

One of the big fears of creeping legalization of marijuana is that rates of use and addiction will soar; it is estimated that nearly 9% of users become dependent⁷. “It’s not the worst drug-use disorder, but it’s also not a good thing — and the relapse rates are very high,” says Margaret Haney, a neurobiologist at Columbia University in New York. She notes that dependence is particularly problematic among younger users, who are less motivated to quit than older people with families and careers.

Some evidence links heavy adolescent use to lasting impairments in cognitive development. Pharmacologist Harold Kalant at the University of Toronto, Canada, cites a 38-year-long study from New Zealand⁸, showing that adolescents who use the drug experience potentially lifelong deficits in cognitive skills, including judgement and working memory. “These kids have much poorer school records, higher dropout rates and poorer employment prospects,” says Kalant. This was the first of only a few studies to investigate this relationship over time, so the results are open to interpretation. “It’s biologically plausible,” says Haney. “But there are no data that definitively demonstrate that marijuana is causing this effect.” Indeed, a study published in August found no difference in mental and physical health among young adults who had used cannabis as adolescents⁹. However, the study was smaller and shorter than the New Zealand one and, although Haney says it was

well conducted, she adds that its size makes it hard to identify real differences between users and non-users. “Those arguing marijuana might cause psychosis would probably not be dissuaded by such small numbers,” she says.

One recent analysis showed that although adolescent use is higher in states that allow medical marijuana, passage of those laws did not lead to increased use — suggesting that pre-existing local attitudes may be the primary factor¹⁰. Benedikt Fischer, a public health researcher at the Centre for Addiction and Mental Health in Toronto, sees the fear that cannabis is a danger to young people as an exaggerated product of ‘Reefer Madness’-style moral panic. “There are a lot of hockey-related brain injuries,” he says, “but nobody in this country would think about prohibiting hockey for young people.” He believes that other risks, such as increased frequency of cannabis-related traffic accidents, may be under-appreciated.

Other critics see a double-standard for cannabis, given the extensive misuse of pharmaceuticals such as opioid painkillers. “There are risks, but those can be modulated by careful patient screening and public health messaging,” says pain specialist Mark Ware, at McGill University in Montreal, Canada, “and they should not stop progress in considering these drugs for middle-aged women with multiple sclerosis, or for men with HIV/AIDS.” **M.E.**

a new generation of researchers. Colorado’s research grant programme supports a placebo-controlled trial for PTSD in 76 veterans and a pain study that pits cannabis against oxycodone. Cannabis growers are also keen to build their scientific case — Canadian company National Green Biomed, based in Vancouver, gave a Can\$1 million (US\$760,000) grant to researchers at the University of British Columbia to explore whether medical marijuana suppresses HIV infection. And Bedrocan has hired its first clinical trial coordinator, and is investigating the chemical composition of its plants in order to learn why patients might claim different benefits from different strains.

The future of medical marijuana remains unclear — the money now pouring into this new industry could accelerate deregulation, or a backlash may lead to stricter controls. From Ware’s perspective, the current situation may offer a fleeting chance for the research community to seize the reins (see page S9). “There’s an opportunity for us to take a global leadership position — the world is crying out for some

direction in managing not just medical, but also recreational cannabis,” he says. “This should stop being an issue for legal minds to wrestle over; it should be the scientists that are putting the evidence together that drive the policy.” ■

Michael Eisenstein is a freelance science writer based in Philadelphia, Pennsylvania.

1. Abrams, D. I. *et al.* *Neurology* **68**, 515–521 (2007).
2. Ware, M. A. *et al.* *Can. Med. Assoc. J.* **182**, E694–E701 (2010).
3. Naftali, T. *et al.* *Clin. Gastroenterol. Hepatol.* **11**, 1276–1280 (2013).
4. Press, C. A., Knupp, K. G. & Chapman, K. E. *Epilepsy Behav.* **45**, 49–52 (2015).
5. Whiting, P. F. *et al.* *J. Am. Med. Assoc.* **313**, 2456–2473 (2015).
6. Hazekamp, A. & Heerdink, E. R. *Eur. J. Clin. Pharmacol.* **69**, 1575–1580 (2013).
7. Volkow, N. D. *et al.* *N. Engl. J. Med.* **370**, 2219–2227 (2014).
8. Meier, M. H. *et al.* *Proc. Natl Acad. Sci. USA* **109**, E2657–E2664 (2012).
9. Bechtold, J., Simpson, T., White, H. R. & Pardini, D. *Psychol. Addict. Behav.* <http://dx.doi.org/10.1037/adb0000103> (2015).
10. Hasin, D. S. *et al.* *Lancet Psychiatry* **2**, 601–608 (2015).

CANNABIS

4 BIG QUESTIONS

As restrictions around cannabis research ease, scientists are exploring how the plant could be medically useful. Here are four of the hardest questions they face.

BY JULIE GOULD

QUESTION

WHAT WE KNOW

WHY IT MATTERS

MAJOR HURDLES

1

How many types of cannabis are there?

Cannabis seems to come in many forms, partly because of millennia of breeding by humans. Producers have allocated a 'folk taxonomy' to strains with various medicinal and chemical properties, but the evolutionary relationship between species is unclear (see page S4).

Researchers need to characterize plants for drug development, including by species or subspecies. Without an accurate taxonomy, there is no comparative base to work from and some varieties might become extinct.

Because cannabis is wind pollinated with separate male and female plants, controlling breeding can be difficult. There is no organized seed bank and legal restrictions impede the collection of samples. Little government funding means there are few research labs.

2

What are the medically useful compounds in cannabis and what diseases can they treat?

Of the 100 or so cannabinoids, tetrahydrocannabinol (THC) and cannabidiol have been the focus of most research. Compounds such as the volatile terpenes also seem to have intriguing properties. All constituents might act alone or in combination.

Those who use medical marijuana claim that the herb can treat a range of ailments, but evidence is thin. To test these claims and create a pharmaceutical product, the active compounds need to be isolated and investigated.

Research has been undermined by a lingering assumption that all cannabinoids share the properties of THC — and regulators apply the same restrictions to all of them (see page S6).

3

Can the endocannabinoid system be targeted by compounds other than those found in cannabis?

So far, drugs that target the endocannabinoid system have failed. Cannabidiol has an anti-schizophrenia effect by increasing levels of the endocannabinoid anandamide, a type of neurotransmitter, suggesting that anandamide is the active molecule.

Therapies based on endogenous molecules are a mainstay of medicine. The hormone insulin was successfully given to patients with diabetes shortly after its discovery. And cortisone-based medicines can be used to relieve pain and reduce swelling.

Endocannabinoids have localized actions and short half-lives, so their use as systemic therapies is challenging. Endocannabinoid receptors are found throughout the body, making it difficult to target specific physiological processes without side effects.

4

What are the best ways to deliver cannabis-based medicines?

Delivery through the lung by smoking or vaporizing facilitates fast absorption. Administration that involves compounds passing through the liver is tricky because this organ destroys cannabinoids. Only an oral, under-tongue spray has so far been approved.

Route of delivery determines how much of a compound is absorbed by different parts of the body, and how quickly. The faster that THC-containing medicines are delivered to the brain, the more potential there is for psychoactive side effects, as well as for abuse.

Regulators require that cannabis-based drugs have no psychoactive effects, so delivery of THC to the brain needs to be controlled. Little is known about the metabolism of other cannabinoids and hence appropriate delivery routes.

Julie Gould is editor of *NatureJobs* in London.

PHARMACOLOGY

Cannabis in neurology—a potted review

Richard Hosking and John Zajicek

Discovery of the endogenous cannabinoid signalling system unleashed substantial new research into several neurological conditions. A recent systematic review suggests that medical marijuana can improve a number of symptoms—particularly spasticity—in multiple sclerosis, but cannabinoids can have adverse psychological effects and their comparative effectiveness is unknown.

Hosking, R. & Zajicek, J. *Nat. Rev. Neurol.* **10**, 429–430 (2014); published online 8 July 2014; doi:10.1038/nrneurol.2014.122

Cannabis attracts attention. The war on drugs and controversy over legalization contrasts with renewed interest in alternative therapies and anecdotal reports of patient benefit. A growing understanding of cannabis pharmacology has revealed important molecular mechanisms that relate to neural communication and inflammation, and several licensed preparations of medical cannabis are now available (Box 1). A recent systematic review published in *Neurology* by Barbara Koppel and colleagues provides timely guidelines for the use of medical marijuana in selected neurological disorders.¹ The evidence summarized in the review by Koppel *et al.* supports a role for cannabis in the symptomatic treatment of multiple sclerosis (MS), but its efficacy in Huntington disease (HD), Parkinson disease (PD) and epilepsy is unknown.

“...cannabinoids should be investigated ... using the same evidence-based criteria as for all other drugs”

Cannabis is thought to derive its name from ancient Sanskrit and Hebrew texts, where it means ‘fragrant cane’, and the plant has been of medicinal interest for millennia. The first clinical study involving cannabis was conducted in 19th-century Calcutta by Irish physician Sir William O’Shaughnessy, who introduced Indian hemp to Victorian medicine.² However, it was not until the 1960s that Raphael

Mechoulam and colleagues in Israel identified the major psychoactive cannabinoid delta-9-tetrahydrocannabinol (Δ 9-THC), which led to the discovery of an extensive endogenous lipid signalling system.³

There are currently over 100 known cannabinoids that have diverse effects at both cannabinoid and non-cannabinoid receptors. Cannabinoid receptor 1 (CB1) is the most common G-protein-coupled receptor in the CNS. High densities of CB1 within the cerebellum, basal ganglia, hippocampus and cerebral cortex correlate with the capacity of cannabis to produce motor and cognitive impairment. This psychoactivity—a property that not all cannabinoids possess—is largely mediated by Δ 9-THC. Cannabinoid receptor 2 (CB2) mRNA is found within cells of the immune system including lymphocytes, monocytes and microglia, although current problems with antibody specificity complicate receptor localization. The major ligands for these receptors are endogenous cannabinoids, such as anandamide and 2-arachidonoylglycerol, which are derived from arachidonic acid and demonstrate marked overlap with prostaglandin signalling.⁴ The actions of these ‘endocannabinoids’ are complex, as are those of plant-derived cannabinoids, which, in addition to receptor stimulation, might have antioxidant properties.

Cannabinoid receptors are present in all major pain pathways, and the analgesic effects of cannabinoids are likely to be mediated by postsynaptic retrograde inhibition of neurotransmission via CB1.

An alternative mechanism might involve CB2 activation on microglia and peripheral inflammatory cells.³ Endocannabinoids are hydrolysed by several enzymes, including fatty acid amide hydrolase (FAAH), monoacylglycerol lipase (MAGL) and diacylglycerol lipase (DAGL). Most information on these mechanisms derives from animal studies, but while animal models have provided many of the important tools for dissecting complex immunological processes, they rarely replicate the pathological features of human disease.⁵

Koppel and colleagues reviewed the full text of 64 papers from a total of 1,730 abstracts published between January 1948 and November 2013.¹ This search yielded 34 studies that met the authors’ inclusion criteria. Eight of the included studies were rated as Class I evidence according to the American Academy of Neurology classification for therapeutic articles. The authors

Box 1 | Cannabinoids

Cannabis sativa

The botanical name for the hemp plant from which cannabis is obtained

Marijuana

A synonym for cannabis that also refers to the dried flowers and leaves

Cannabinoids

Plant-derived compounds and endogenous or synthetic analogues

Delta-9-tetrahydrocannabinol (Δ 9-THC)

The main plant-derived psychoactive cannabinoid; the relative amount of Δ 9-THC in a preparation determines the extent of psychological effects

Cannabidiol (CBD)

The main non-psychoactive plant-derived cannabinoid

Endocannabinoids

Endogenous lipid signalling molecules that are widespread throughout the body

Oral cannabis extract

Extracts such as Cannador, a capsule with a defined THC:CBD ratio (2.50:1.25 mg)

Dronabinol (Marinol)

Oral synthetic Δ 9-THC

Nabilone (Cesamet)

Oral synthetic cannabinoid similar to Δ 9-THC

Nabiximols (Sativex)

Herbal oromucosal spray with a defined THC:CBD ratio (2.7:2.5 mg)

addressed several questions, including the efficacy of cannabinoids for treating MS symptoms of spasticity, central pain, painful spasms, bladder dysfunction and involuntary movements including tremor. They also assessed the use of cannabinoids in HD, levodopa-induced dyskinesias in PD, cervical dystonia, the tics of Tourette syndrome, and seizure frequency in epilepsy.

The main finding of Koppel and co-workers was that oral cannabis extract (see Box 1) is effective in treating spasticity, central pain and painful spasms in patients with MS. Nabiximols and THC are also likely to be effective for these indications. Nabiximols is probably effective in reducing urinary frequency, but no formulation was felt to improve MS-related tremor. Oral cannabis extract was found to be probably ineffective in treating levodopa-induced dyskinesias in PD; no discernible effect was found for any cannabinoid on the remaining neurological disorders. Koppel *et al.* point out that the risk of serious adverse psychological effects was nearly 1%. The comparative effectiveness of medical marijuana relative to other therapies for these indications is unknown.

We welcome evidence-based guidelines for cannabinoid use in neurological disease, and Koppel *et al.* have provided an excellent systematic review of current clinical research. The limitations of the present study reflect those of the original studies. These include problems with subjective patient-centred rating scales; objective clinical scales with potential interobserver variability; the risk of unblinding owing to the psychoactive properties of the drugs; and often-substantial placebo effects that make treatment effects difficult to establish.

An additional issue, not specific to cannabis preparations, is the difficulty of identifying responders within heterogeneous populations, which has led to novel trial designs and will continue to tax medical ingenuity. Importantly, the studies included in the Koppel *et al.* review used a wide range of formulations with different cannabinoid content, methods of administration and dose. Cannabinoids are intensely lipophilic and readily cross the blood–brain barrier, but the absorption and bioavailability of preparations are highly variable, partly due to first-pass metabolism (metabolization before the drug reaches the systemic circulation), which oromucosal administration attempts to reduce.

Titration speed can affect the development of unwanted psychoactive effects, but the current available evidence suggests that these effects resolve on dose reduction or drug cessation. Smoked marijuana causes the most rapid rise in plasma THC concentration, which then quickly falls as a result of tissue distribution.³ However, combustion alters cannabinoid activity, and the attendant risks of inhaled carcinogens are an additional cause for concern. Furthermore, a recent study shows that patients with MS who smoke cannabis have greater cognitive impairment than those who do not use this drug.⁶

Beyond symptom amelioration, the potential of cannabinoids as neuroprotective agents has created a great deal of interest. Although the overall results of a recent dronabinol study of neuroprotection in progressive MS were negative, subgroup analysis did suggest a possible early treatment effect in less-disabled patients.⁷ Polymorphisms in the gene encoding CB1 could influence the inflammatory neurodegenerative process in MS,⁸ and receptor downregulation after continued exposure to THC might mitigate the effects of this drug in certain brain regions.⁹ In patients with HD, loss of CB1 from the striatum could be an important pathogenic factor. Data from animal models suggest that abnormal huntingtin protein indirectly reduces CB1 expression in striatal neurons, leading to increased excitotoxicity and decreased levels of brain-derived neurotrophic factor. Early treatment with CB1 agonists might prevent these changes.¹⁰

Cannabidiol lacks psychoactivity and has shown promise as a neuroprotective agent in preclinical studies. An alternative strategy to prevent adverse psychological reactions that could also be of use in neuroprotective studies is to increase endocannabinoid signalling or ‘tone’ using inhibitors of the constitutive hydrolytic enzymes FAAH, MAGL and DAGL. In fact, this could be one mechanism by which cannabidiol acts, because its general activity at cannabinoid receptors is low.

Cannabis research has paralleled advances in opioid pharmacology, whereby a psychoactive plant extract has led to the discovery of endogenous signalling systems with therapeutic relevance. We agree that cannabinoids should be investigated and prescribed using the same evidence-based criteria as

for all other drugs. Systematic assessment of this evidence, such as the present review by Koppel and co-workers, can greatly aid clinical decisions. The current data, however, are limited, and we look forward to future randomized controlled trials investigating the symptom-alleviating and neuroprotective potential of cannabinoids.

Clinical Neurology Research Group, Plymouth University Peninsula Schools of Medicine and Dentistry, Room N13, ITTC Building 1, Davy Road, Tamar Science Park, Plymouth PL6 8BX, UK (R.H., J.Z.).

Correspondence to: J.Z.

john.zajicek@plymouth.ac.uk

Acknowledgements

J.Z. has received funding to conduct studies of cannabinoids in multiple sclerosis from the UK Medical Research Council.

Competing interests

J.Z. has received funding from the Institut für klinische Forschung, Berlin, which manufactures Cannador. R.H. declares no competing interests.

1. Koppel, B. S. *et al.* Systematic review: efficacy and safety of medical marijuana in selected neurologic disorders: report of the Guideline Development Subcommittee of the American Academy of Neurology. *Neurology* **82**, 1556–1563 (2014).
2. O'Shaughnessy, W. B. On the preparations of the Indian Hemp, or Gunjah, (*Cannabis Indica*): their effects on the animal system in health, and their utility in the treatment of tetanus and other convulsive diseases. *Prov. Med. J. Retrospect. Med. Sci.* **123**, 363–369 (1843).
3. Hosking, R. D. & Zajicek, J. P. Therapeutic potential of cannabis in pain medicine. *Br. J. Anaesth.* **101**, 59–68 (2008).
4. Nomura, D. K. *et al.* Endocannabinoid hydrolysis generates brain prostaglandins that promote neuroinflammation. *Science* **334**, 809–813 (2011).
5. Mestas, J. & Hughes, C. C. Of mice and not men: differences between mouse and human immunology. *J. Immunol.* **172**, 2731–2738 (2004).
6. Pavisian, B. *et al.* Effects of cannabis on cognition in patients with MS: a psychometric and MRI study. *Neurology* **82**, 1879–1887 (2014).
7. Zajicek, J. *et al.* Effect of dronabinol on progression in progressive multiple sclerosis (CUPID): a randomised, placebo-controlled trial. *Lancet Neurol.* **12**, 857–865 (2013).
8. Rossi, S. *et al.* Association between a genetic variant of type-1 cannabinoid receptor and inflammatory neurodegeneration in multiple sclerosis. *PLoS One* **8**, e82848 (2013).
9. Lazenka, M. F., Selley, D. E. & Sim-Selley, L. J. Brain regional differences in CB1 receptor adaptation and regulation of transcription. *Life Sci.* **92**, 446–452 (2013).
10. Blazquez, C. *et al.* Loss of striatal type 1 cannabinoid receptors is a key pathogenic factor in Huntington's disease. *Brain* **134**, 119–136 (2011).

Early phytocannabinoid chemistry to endocannabinoids and beyond

Raphael Mechoulam, Lumír O. Hanuš, Roger Pertwee and Allyn C. Howlett

Abstract | Isolation and structure elucidation of most of the major cannabinoid constituents — including Δ^9 -tetrahydrocannabinol (Δ^9 -THC), which is the principal psychoactive molecule in *Cannabis sativa* — was achieved in the 1960s and 1970s. It was followed by the identification of two cannabinoid receptors in the 1980s and the early 1990s and by the identification of the endocannabinoids shortly thereafter. There have since been considerable advances in our understanding of the endocannabinoid system and its function in the brain, which reveal potential therapeutic targets for a wide range of brain disorders.

The plant *Cannabis sativa* and its many preparations (for example, marijuana, hashish, bhang and ganja) have been used for millennia for recreation (and at times for the achievement of religious ecstasy) as well as in medicine. In ancient China, cannabis was prescribed (together with other plants, as is customary in Chinese medicine) for numerous diseases, but it was noted that when taken in excess it could lead to ‘seeing devils’. In Assyria (about 800 BC), it was named both gan-zi-gun-nu (‘the drug that takes away the mind’) and azallu (when used as a therapeutic). In India, ancient Persia and medieval Arab societies, cannabis use proceeded along these two divergent routes¹. In many countries, hemp — a strain of *Cannabis sativa* that does not cause psychoactivity — was grown for its durable fibres. Our present-day society follows a long tradition of recreational, industrial and medical cannabis use.

Cannabinoid discovery — early history

The behavioural effects of cannabis, in several animal species as well as in humans, were observed in the mid-nineteenth century² (FIG. 1). These experimental observations led to the first attempts to isolate the active constituents of the plant, as had already been done with other plants that

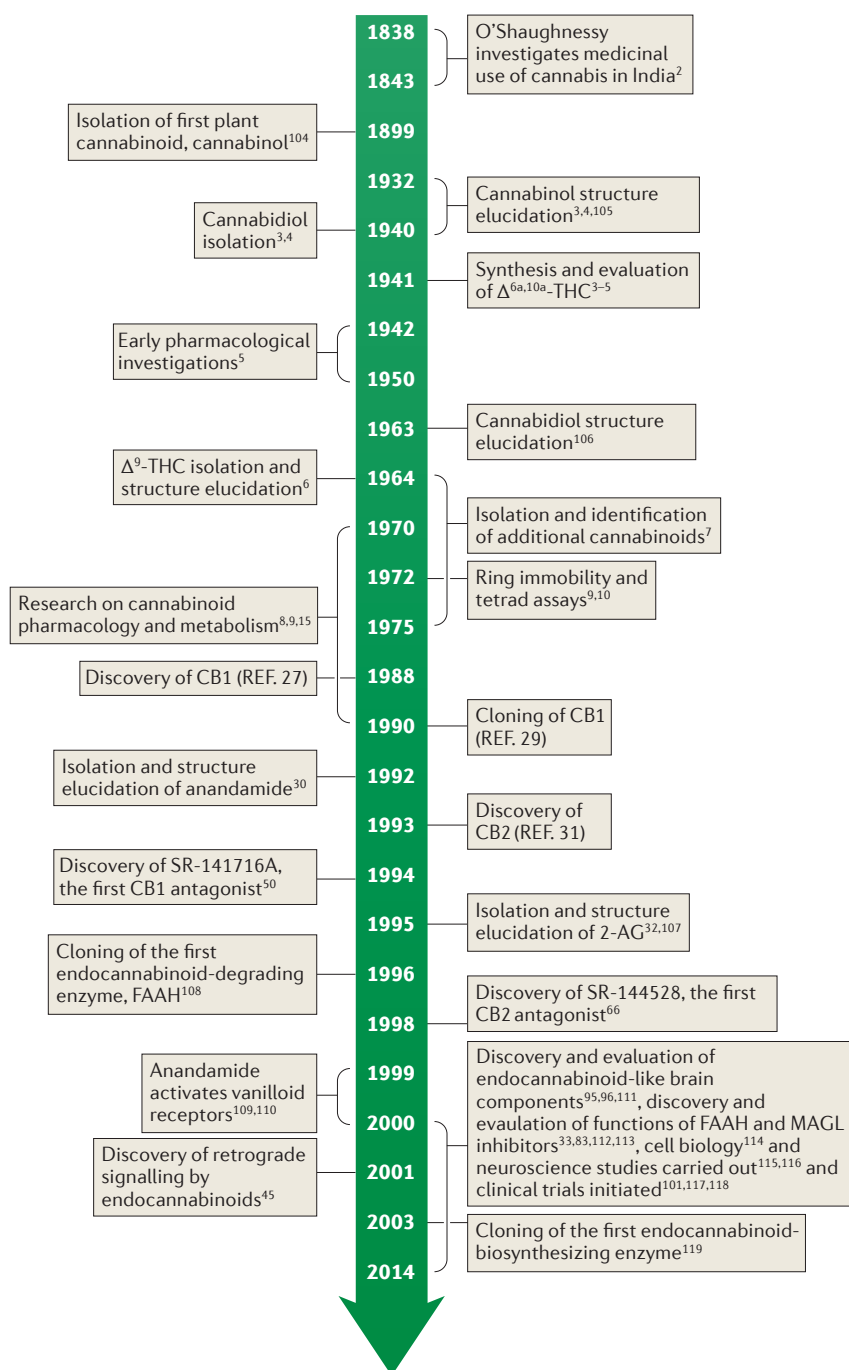
had known neuropharmacological activity — for example, the isolation of morphine. A prize was even awarded in 1855 for the ‘successful’ accomplishment of this project. However, the first isolation of a plant cannabinoid — named cannabinol (CBN) — was not achieved until the end of the nineteenth century. Its structure was elucidated much later, in the 1930s, by the groups of Cahn and Todd in the United Kingdom and by Adams in the United States, when a further component, cannabidiol (CBD), was isolated; however, its structure could not be elucidated at that time. Although considerable effort was invested on the isolation and the elucidation of the structure of the main psychoactive constituents of cannabis, this goal was not reached at that time. A synthetic compound, $\Delta^{6a,10a}$ -tetrahydrocannabinol ($\Delta^{6a,10a}$ -THC), showed pharmacological activity that paralleled the activity of cannabis extracts. Therefore, it was assumed that $\Delta^{6a,10a}$ -THC was chemically related to the active compounds of the plant (FIG. 2). Much of the early research in this area was done using synthetic $\Delta^{6a,10a}$ -THC, which is now known to be considerably less potent than the actual natural product. The chemical and pharmacological work that was carried out until the mid 1940s has been reviewed

elsewhere^{3–5}. Some $\Delta^{6a,10a}$ -THC analogues were even tested in humans. In light of recent media reports about the action of cannabinoids in paediatric epilepsy, it is of interest to note that a derivative of synthetic $\Delta^{6a,10a}$ -THC (at doses of 1.2–1.8 mg daily) was administered to a small number of children with epilepsy and showed positive results. Historical cannabis use in medicine over the ages and early chemical investigations are reviewed in REF. 1.

The reasons for the lack of progress were mostly technical. We now know that cannabinoids are present in cannabis as a mixture of many closely related constituents — over 100 — which were difficult to separate using the methods that were available in the nineteenth and early twentieth centuries. As the active constituents of cannabis were not available in pure form, there was very little biological or clinical work done in this area from the late 1940s until the mid 1960s.

By the 1960s, chromatography methods were well developed for the isolation of pure compounds from mixtures and the availability of novel spectrometric methods meant that the elucidation of the structure of these compounds was possible. Indeed, many cannabinoids were isolated, including Δ^9 -THC, which was reported by Gaoni and Mechoulam in 1964 (REF. 6) (FIG. 2). Their structures were mainly elucidated using NMR, which was a modern method at the time. Several total syntheses of these compounds have been reported and most cannabinoids are now available as both natural and synthetic products. The chemical work until the mid 1970s is reviewed in REF. 7.

The next step in cannabinoid research was the elucidation of the metabolism of Δ^9 -THC and later of CBD. The major metabolic pathway of Δ^9 -THC is hydroxylation, which leads to the formation of an active metabolite, followed by its further oxidation to an inactive acid, which then binds to a sugar molecule. These acid-derived metabolites are stored in fatty tissues and are slowly released⁸. Indeed, the major final Δ^9 -THC metabolite (a carboxylic acid that is present as a glucuronide) can be detected in human urine for several weeks after cannabis use (FIG. 2).



Early neuropharmacology

The advances in the chemistry of plant and synthetic cannabinoids led to renewed interest in their neuropharmacology. Loewe⁵ had found that cannabis extracts (presumably containing high levels of what is now known to be Δ^9 -THC and additional phytocannabinoids) can induce catalepsy in mice and that CBN can also produce this effect, albeit much less potently than the impure THC isolated from the resin. It was these findings that prompted the development by Pertwee⁹ in 1972 of a quantitative *in vivo* assay for psychotropic cannabinoids, known as the ring test, in which the proportion of time that a mouse placed across an elevated horizontal ring remains immobile or cataleptic is measured over a 5 minute period⁹. Martin¹⁰ later used this assay, along with three other bioassays, in what came to be known as the 'mouse tetrad assay'¹⁰. These other assays provide measures of cannabinoid-induced hypokinesia, hypothermia and antinociception in mice, using a tail flick or hot plate test. The mouse tetrad assay is a useful *in vivo* screen for psychotropic cannabinoids, all of which, in contrast to many other types of drugs, generally show similar potency in all four of these bioassays. It was also discovered in the 1940s that cannabinoids can elicit central excitant activity in rabbits and mice and corneal areflexia in rabbits, and that some phytocannabinoids, particularly CBD, can prolong barbiturate-induced sleep by a mechanism that was subsequently discovered to involve the inhibition of certain cytochrome P450 (CYP) enzymes¹⁰.

Following its identification as the main psychoactive constituent of cannabis, Δ^9 -THC attracted particular attention^{10,11}; for example, results obtained from several investigations on humans indicated that when Δ^9 -THC was taken orally or intravenously or when it was inhaled in smoke, it showed substantial potency at producing psychological changes similar to those reportedly experienced in response to recreationally consumed cannabis¹¹. A few other phytocannabinoids, such as CBN, were found to induce cannabis-like effects in humans with low potency (an exception being Δ^8 -THC but there is usually very little Δ^8 -THC in cannabis)¹¹.

It is noteworthy that one synthetic analogue of Δ^9 -THC, nabilone (Cesamet; Valeant Pharmaceuticals North America) was approved in 1981 as a medicine for the suppression of the nausea and vomiting that is produced by chemotherapy¹². Synthetic Δ^9 -THC, dronabinol (Marinol; Solvay Pharmaceuticals, Inc) subsequently entered

Figure 1 | Cannabinoid and endocannabinoid research — a timeline. Almost all early research was devoted to clarification of cannabinoid chemistry^{3,4,104,105}, and pharmacology was mainly done using synthetic compounds⁵. Following the isolation and structure elucidation of the plant cannabinoids, particularly of cannabidiol¹⁰⁶ and of Δ^9 -tetrahydrocannabinol (Δ^9 -THC)⁶, pharmacological and physiological work was initiated^{8,9,15}. The identification of cannabinoid receptors^{24,29,31}, of endogenous cannabinoids^{30,32,107} and of receptor antagonists^{50,66} made possible extensive pharmacological and neurobiological research leading to cloning of the anandamide-degrading enzyme fatty acid amide hydrolase (FAAH)¹⁰⁸, the discovery of retrograde signaling by 2-arachidonoyl glycerol (2-AG)⁴⁵, the discovery of allosteric sites on cannabinoid receptor 1 (CB1)³³, the discovery that endocannabinoids bind to receptors other than CB1 and CB2 (REFS 109–111), the discovery and evaluation of endocannabinoid-like molecules in the brain^{95,96} and the discovery and function of inhibitors of the endocannabinoid-degrading enzymes^{112,113}. Cell biology¹¹⁴ and neuroscience^{115,116} investigations were also carried out, and clinical trials were initiated^{101,117,118}. Cloning of DAG lipase was also reported¹¹⁹.

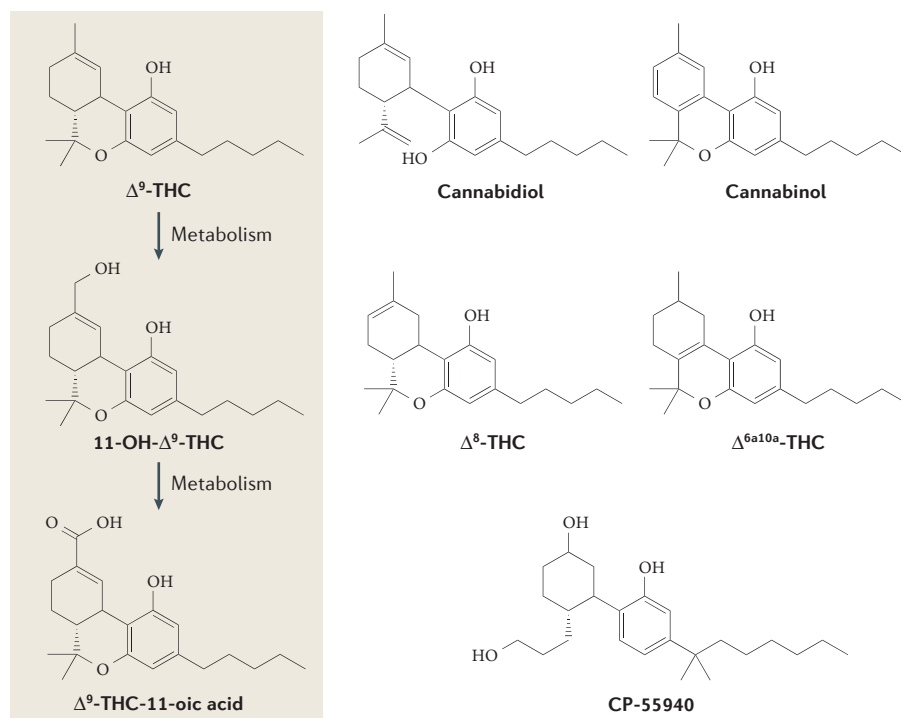


Figure 2 | A major metabolic pathway of Δ^9 -THC and the structures of some plant and synthetic cannabinoids. **a** | The major psychoactive cannabis constituent, Δ^9 -tetrahydrocannabinol (Δ^9 -THC), is first metabolized by enzymatic hydroxylation to produce psychoactive 11-hydroxy- Δ^9 -THC (11-OH- Δ^9 -THC) and then by enzymatic oxidation to non-psychoactive Δ^9 -THC-11-oic acid, which is stored in fatty tissues as a glucuronide and is slowly released. The glucuronide may be detected in the urine for several weeks after a single cannabis use. **b** | The structures of some plant and synthetic cannabinoids. Δ^9 -THC, the plant constituents cannabinol and Δ^8 -THC, and synthetic $\Delta^{6a,10a}$ -THC and CP-55940 cause cannabis-type psychoactivity, whereas cannabidiol does not.

the clinic as a licensed medicine, in 1985 as an antiemetic and in 1992 as an appetite stimulant¹². Claims from patients that cannabis can ameliorate unwanted symptoms of multiple sclerosis also encouraged the development of the cannabis-based medicine naviximols¹² (Sativex; GWPharma), which contains both Δ^9 -THC and the non-psychoactive CBD; this was first licensed as a medicine in 2005 in Canada for the relief of pain experienced by adult patients suffering from multiple sclerosis or advanced cancer, and subsequently as a medicine to ameliorate spasticity caused by multiple sclerosis¹².

Discovery of the cannabinoid receptors

Although a considerable amount of pharmacological work was done on the activity of Δ^9 -THC, its mechanism of action was not elucidated for more than 20 years after its identification. Indeed, it was originally thought that the mode of action of Δ^9 -THC was nonspecific in nature and that it might involve interactions with lipid membranes. However, although the stereospecificity of the action of Δ^9 -THC and related synthetic

cannabinoids^{13,14}, as well as pharmacological studies, in humans and animals had suggested a putative cannabinoid receptor^{15,16}, it was not until the 1980s that evidence for a protein receptor was sought.

As the family of known G proteins expanded in the late 1970s and early 1980s, so did the list of receptors for hormones and neurotransmitters to which they could couple. Agonists of opioid, muscarinic, cholinergic and α -adrenergic receptors resulted in inhibition of G_s -stimulated adenylyl cyclase^{17–19}, and functional homology with these neuromodulators led to the discovery that cannabinoids also inhibited this enzyme^{20–22} by a pertussis toxin-sensitive mechanism²³. This clearly indicated that the cannabinoid receptor was a G protein-coupled receptor (GPCR).

From the structure–activity relationship (SAR) established using cannabimimetic compounds from Pfizer Central Research, the Howlett laboratory identified CP55940 (FIG. 2) as a highly potent cannabinoid analogue and, in 1988, reported the determination and characterization of a cannabinoid receptor from the brain for which the criteria for a

high-affinity, stereoselective receptor in brain tissue had been fulfilled²⁴. Competitive displacement of [³H]CP55940 from its target in rat brain membranes by cannabinoid agonists was enantioselective and followed the order of potency for both G_i -mediated inhibition of adenylyl cyclase as well as antinociception in several rodent models^{24–27}. Later, signal transduction assays were used to ultimately deorphanize a 7-transmembrane receptor now known to be the cannabinoid receptor 1 (CB1; also known as CNR1)^{28,29}.

Discovery of endocannabinoids and CB2

Receptors are mostly activated by endogenous molecules, and therefore, there was a strong reason to look for endogenous cannabinoids. As Δ^9 -THC and its related compounds that bind to the CB1 are lipids, it was reasonable to assume that any endogenous cannabinoids would also be lipids. In order to isolate putative endogenous cannabinoid compounds, the ability of porcine brain extracts to displace a novel, highly potent radiolabelled cannabinoid probe, [³H]HU-243, bound to CB1 was tested in the Mechoulam laboratory. The fractions that inhibited the binding of [³H]HU-243 to the cannabinoid receptor were purified by a series of chromatographies, which ultimately led to the generation of a minute amount of a single compound, an amide of arachidonic acid — arachidonoyl ethanolamide — which was named anandamide; this was the first endocannabinoid to be identified³⁰. The structure of anandamide (FIG. 3) was established by mass spectrometry, NMR spectroscopy and by its synthesis³⁰. Anandamide was found to have inhibitory activity that was equivalent to that of Δ^9 -THC and was subsequently shown to have cannabimimetic activity as it inhibited the twitch response of isolated mouse vasa deferentia³⁰.

In the meantime, a second receptor, CB2 (also known as CNR2), had been identified by sequence homology³¹ and was presumed to be mainly present in the periphery; therefore, a search for a ‘peripheral’ endogenous agonist was initiated. Using the same techniques that were used to isolate anandamide, it was possible to isolate an ester of arachidonic acid — 2-arachidonoyl glycerol (2-AG)³² — from canine intestines (FIG. 1). This compound was unexpectedly found to bind CB1 and CB2 and to inhibit adenylyl cyclase with a potency similar to that of Δ^9 -THC. 2-AG also shared the ability of Δ^9 -THC and anandamide to inhibit electrically evoked contractions of isolated mouse vasa deferentia; however, it was less potent than Δ^9 -THC³². Following administration

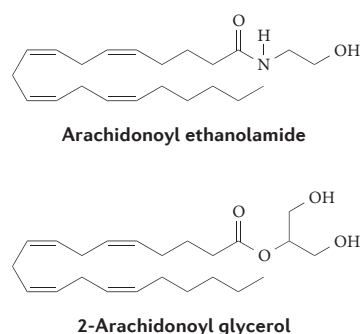


Figure 3 | Structures of the main endocannabinoids, anandamide and 2-AG, which bind to CB1 and CB2 endocannabinoid receptors. Arachidonoyl ethanolamide (also known as anandamide) and 2-arachidonoyl glycerol (2-AG) are hydrolysed to arachidonic acid by the enzymes fatty acid amide hydrolase (FAAH) and monoacylglycerol lipase (MAGL), respectively. Blocking these enzymes with various synthetic compounds leads to increased levels of these endocannabinoids.

to mice, both anandamide and 2-AG caused the typical tetrad of effects produced by Δ^9 -THC: antinociception, immobility, reduction of spontaneous activity and lowering of rectal temperature. Although a few additional endocannabinoids have been reported, none of them has been confirmed as a natural endocannabinoid.

Anandamide is a partial agonist for CB1 and CB2 and shows less relative intrinsic activity (also known as relative intrinsic efficacy) and affinity for CB2 than for CB1. 2-AG shows greater potency and efficacy than anandamide as a CB1 agonist and greater potency than anandamide as a CB2 agonist³³. In addition, it has been found that both endocannabinoids interact with certain non-CB1 and non-CB2 receptors and ion channels³³. In the past few years, lipoxin A4 and a new family of peptides (known as pepcans) have been reported to target CB1 as allosteric modulators^{34,35} and the peptide hemopressin, which is a putative brain constituent, has been found to lower pain via action on a cannabinoid receptor³⁶.

Synthesis of cannabinoid analogues that have high affinity and specificity for CB2 was achieved in the mid to late 1990s^{37,38} and led to the discovery of the role of CB2 in immunosuppression, neuroprotection and neuropathic and inflammatory pain. This consequently led to considerable interest in developing and investigating CB2-selective agonists^{39–43}.

Both anandamide and 2-AG are synthesized on demand, often in response to increased concentration of intracellular

calcium⁴⁴, and it is now generally accepted that one important role of these endocannabinoids, although possibly only of 2-AG, is to function as retrograde synaptic messengers that can prevent the development of excessive neuronal activity in the central nervous system and thereby contribute to the maintenance of homeostasis in both health and disease⁴⁵. Thus, there is good evidence that neurotransmitters, such as glutamate, produce postsynaptic increases in the concentration of intracellular calcium in a manner that can induce postsynaptic biosynthesis and release of anandamide or 2-AG into the synapse. In turn, this induces subsequent endocannabinoid-induced activation of presynaptic CB1, which causes an inhibition of the neuronal release of glutamate, γ -aminobutyric acid or other neurotransmitters in brain areas that include the cerebral cortex, hippocampus, ventral tegmental area, substantia nigra, hypothalamus and cerebellum^{46–48}. There is also evidence that, when produced postsynaptically in response to the activation of postsynaptic metabotropic glutamate receptor 5 (MGLUR5), anandamide activates postsynaptic transient receptor potential cation channel subfamily V member 1 (TRPV1) channels⁴⁸. It is also noteworthy that results obtained from *in vivo* experiments with rats suggest that retrograde 2-AG signalling that is triggered by the activation of MGLUR5 can suppress pain sensitivity⁴⁹. The endocannabinoid retrograde transport mechanism and modulation of synaptic transmission have not yet been fully elucidated^{46–48}.

Search for antagonist ligands

The holy grail for cannabinoid synthetic chemists was an antagonist that could block the effects of Δ^9 -THC. It seems quite unusual that no natural product or structurally related analogue emerged to block the cannabinoid receptors. Before the advent of gene knockout techniques, it was difficult to establish whether a pharmacological effect was mediated by a receptor if a selective antagonist for that receptor had not been developed. Thus, one can imagine the excitement generated at an International Cannabinoid Research Society meeting in 1993 when a team of researchers from the French pharmaceutical company Sanofi Recherche announced their discovery of an antagonist for CB1, SR141716A⁵⁰. This compound was radiolabelled to investigate receptor pharmacology⁵¹ and was soon modified to develop the first ligands for *in vivo* imaging⁵². The discovery of an antagonist (SR141716A), which was in fact subsequently identified as an inverse agonist, helped to characterize

additional cellular signalling pathways for CB1 (REFS 50, 51, 53–55). More importantly, an antagonist could finally be used to identify animal behaviours that were truly due to CB1 activation^{56–58}. Indeed, the syndrome of 'dependence' on cannabinoid agonists was first shown in an animal model after precipitated withdrawal using SR141716A^{59,60}. Within a short period of time, industrial laboratories and academic research groups reported the synthesis of additional CB1 antagonists and inverse agonists^{61–64}.

The first CB2-selective antagonists AM630 (also known as iodopravodoline) and SR144528 emerged in the mid 1990s^{65,66} and increased the ability to discern novel actions that could be attributed to CB2, including actions observed in liver Kupfer cells⁶⁷, microglial cells and astrocytes^{68,69} and in the gastrointestinal system⁷⁰, among others. Since that time, there has been considerable progress towards the development of highly selective and potent CB2 antagonists^{41,71}.

SR141716A (also known as rimonabant) is used therapeutically for the treatment of obesity-related metabolic syndrome components, including dyslipidaemia and diabetes^{72–74}. SR141716A was marketed in Europe but failed to gain approval from the US Food and Drug Administration. As might be predicted, a drug that blocks CB1 neuromodulation at synapses for the major stimulatory (in the case of glutamate) and inhibitory (in the case of GABA) transmitters throughout the brain would be likely to produce multiple 'off-target' effects. One such side effect, which was reported in 2009, was an increase in reported signs of depression in vulnerable individuals treated with SR141716A^{75,76}. It could be argued that the benefit to risk ratio in a morbidly obese patient population might mitigate the concerns about depression. However, the drug was withdrawn from the market and similar analogues from other pharmaceutical companies were taken out of the development pipeline. Nevertheless, the development of SR141716A by Sanofi–Aventis can be considered to be a major contributor to our understanding that CB1 is present and functional in tissues such as adipose, liver and pancreas under pathological conditions of high-fat diet or obesity⁷⁷. This new understanding of the role of CB1 in metabolic regulation has inspired the search for novel antagonists that fail to gain access to the brain^{78,79}. An alternative clinical strategy would be to screen for individuals who might be most susceptible to the limbic effects of CB1 antagonists before selecting a treatment modality⁸⁰.

Endocannabinoid neuropharmacology

The discovery that anandamide and 2-AG are endocannabinoids prompted research to identify the biochemical processes that are responsible for both their biosynthesis and their metabolism. This research showed that these two endocannabinoids are synthesized 'on demand' rather than stored, and it identified biosynthetic and metabolic pathways for both of them^{81–83}. Thus, it has been discovered that 2-AG is formed from diacylglycerol (DAG) in a process that is catalysed by *sn*1-specific DAG lipase- α and lipase- β , and that the main biosynthetic pathway for anandamide involves the formation of *N*-arachidonoyl phosphatidylethanolamine (NAPE) from phosphatidylethanolamine and phosphatidylcholine, which is catalysed by an as yet uncharacterized calcium-dependent transacylase enzyme. This is then followed by the conversion of NAPE to anandamide in a single step that is catalysed by NAPE-selective phospholipase D and/or in two or three steps that are catalysed by other enzymes. It has also been found

that, following their release, anandamide and 2-AG are mainly metabolized to arachidonic acid, the major metabolizing enzymes being fatty acid amide hydrolase (FAAH) for anandamide and monoacylglycerol lipase (MAGL) for 2-AG^{81,82}. Other endocannabinoid-metabolizing enzymes include FAAH-2 for anandamide, α , β -hydrolase domain-containing 6 (ABDH6) and ABDH12 for 2-AG, and cytochrome P450 enzymes, lipoxygenases and cyclooxygenase 2 for both of these endocannabinoids^{81,82}. The physiological relevance of the lipoxygenase and cyclooxygenase derivatives of anandamide and 2-AG is not yet clear. It is also noteworthy that anandamide and 2-AG can undergo cellular uptake following their release, although whether this process is mediated by a transporter is currently unclear^{81,82}.

It is now recognized that, although engineering exogenous cannabinoids provided insights into receptor usage and linked functional events, the intracellular and extracellular actions and fate of endocannabinoids

versus those of exogenously introduced cannabinoids may differ and have different physiological consequences^{33,44}. It is also recognized that many cannabinoid receptor ligands also interact with a wide range of non-cannabinoid receptor targets and that, irrespective of whether they are endogenous, synthetic or plant cannabinoids, the pharmacological profiles of these compounds often vary considerably from each other^{33,44}.

The endocannabinoid receptors, the endocannabinoids and their biosynthetic and biodegrading enzymes constitute what has come to be known as the endocannabinoid system, the discovery of which prompted a search for its physiological and pathophysiological roles. This search revealed that there are several disorders in which endocannabinoids are released to their receptors in an 'autoprotective' manner that ameliorates unwanted effects of these disorders^{82–84}. It also raised the possibility that increasing extracellular levels of a released endocannabinoid by inhibiting metabolizing enzymes such as FAAH or MAGL, or by inhibiting the cellular uptake of anandamide, might prove to be an effective therapeutic strategy to manage some of these disorders, which include multiple sclerosis, Parkinson's disease, schizophrenia, hypertension, inflammatory bowel diseases, pruritus, Alzheimer's disease, depression, obsessive compulsive disorder and cancer^{82–84}.

The discovery of the endocannabinoid system also led to a reinvestigation of the interactions of plant and synthetic cannabinoids with this system and other biochemical entities. As a result, evidence has emerged that Δ^9 -THC targets receptors other than CB1 (REFS 85–87). For example, at submicromolar concentrations, Δ^9 -THC has also been found to have several effects: first, it has been found to activate CB2, albeit with less efficacy than it activates CB1 (REF. 88); second, it has been found to activate the orphanized GPCRs GPR18 (REF. 89) and GPR55 (REF. 33), the cation channels TRPA1 and TRPV2 (REFS 90) and the nuclear receptor peroxisome proliferator-activated receptor- γ (PPAR γ)³³; third, it has been found to block the activation both of 5-hydroxytryptamine 3 (5HT3) ligand-gated ion channels^{33,85} and of TRPM8 cation channels⁹⁰; and, last, it has been found to enhance the activation both of α 1 subunits and α 1 β 1 dimers of human glycine ligand-gated ion channels and of native glycine receptors in rat isolated ventral tegmental area neurons³³. There have also been reports that submicromolar concentrations of Δ^9 -THC can inhibit the enzyme lysophosphatidylcholine acyl transferase¹¹, that it can increase the activity

Glossary

Affinity

The potency with which a compound binds to a particular receptor; the higher the affinity of the compound, the lower the concentration at which it achieves a given level of receptor occupancy.

Agonists

Compounds that can activate pharmacological receptors; a full agonist is more potent than a partial agonist and so usually produces a greater maximum functional response.

Allosteric modulators

Drugs that can act on an allosteric site of a receptor to increase or to reduce the ability of an agonist or an inverse agonist to induce a functional response when it targets a different (orthosteric) site on the same receptor.

Antagonist

A compound that can bind to, but cannot activate, a receptor by targeting its orthosteric site and that can therefore prevent both drug-induced agonism and drug-induced inverse agonism at this receptor.

Antinociception

Another term for pain relief.

Apoptosis

A process of programmed cell death that usually has advantageous consequences.

Catalepsy

A condition that is characterized by immobility and muscular rigidity.

Endocannabinoid

An endogenous compound that can directly activate or block cannabinoid CB1 and/or CB2 or that can act as a positive or negative allosteric modulator to increase or to reduce responses of CB1 and/or CB2 to direct agonists or inverse agonists.

G protein-coupled receptor

(GPCR). A seven-transmembrane domain receptor that induces G-protein-mediated activation of intracellular signal transduction pathways when occupied by an agonist.

Hashish

A cannabis-derived preparation that consists mostly of dried cannabis resin.

Hypokinesia

A condition that is characterized by decreased bodily movement.

Inverse agonist

A compound that binds to a receptor in a manner that induces a pharmacological response opposite to the response that is induced by an agonist for the same receptor.

Relative intrinsic activity

The relative ability of drug–receptor complexes to produce maximum functional responses; a high-efficacy agonist needs to occupy fewer receptors to produce a maximal response than a low-efficacy agonist (also known as a partial agonist).

Retrograde synaptic messengers

Compounds that are released by a postsynaptic dendrite or cell body, but that act presynaptically — for example, to influence the release of a transmitter.

Structure–activity relationship

(SAR). The relationship between the pharmacological activity of compounds and their chemical structures.

Transient receptor potential cation channel subfamily V member 1

(TRPV1). A member of a superfamily of transmembrane cation channels; it was previously known as vanilloid receptor 1.

of phospholipase C, which can catalyse the production of DAG and phospholipase A2 (REF. 11) and that it can both inhibit the uptake of adenosine by cultured microglia and macrophages and affect the synaptosomal uptake of 5-hydroxytryptamine (it inhibits this process), of noradrenaline (it enhances this process) and of dopamine (it both enhances and inhibits this process)^{85,87}. In addition, at higher concentrations, Δ^9 -THC has been found to affect several other such pharmacological targets^{85,87}. For example, at concentrations between 1 μ M and 10 μ M, it has been reported to enhance the activation of β -adrenoceptors, to function as a negative allosteric modulator of μ - and δ -opioid receptors, to activate the cation channels TRPV3 and TRPV4 and to inhibit T-type calcium (Ca_v3) and potassium ($\text{K}_v1.2$) voltage-gated ion channels, as well as conductance in Na^+ voltage-gated ion channels. In this concentration range, Δ^9 -THC has also been reported to inhibit the enzymes lipoxygenase, Na^+/K^+ -ATPase and monoamine oxidase, as well as the cytochrome P450 enzymes CYP1A1, CYP1A2, CYP2B6 and CYP2C9, to inhibit noradrenaline-induced melatonin biosynthesis, and to activate or to inhibit Mg^{2+} -ATPase^{85,87}.

Perspectives

There has been much progress in our understanding of the plant cannabinoids and of CB1 and CB2. We have identified endogenous lipid mediators that act on these receptors to regulate multiple pathways of cellular signalling. We have discovered synthetic agonists and antagonists for these receptors as well as allosteric modulators of CB1. However, there is still much more knowledge to be gained and challenges to be met in the fields of cannabinoid receptor neuroscience, pharmacology, molecular biology and cannabinoid medicine.

We now need to understand how the endocannabinoid receptors interact with other proteins in complexes that regulate differentiated functions both at the cell surface and in intracellular organelles, particularly in the brain^{91–93}.

Dozens of endogenous molecules, with structures that resemble those of the endocannabinoids, have been discovered in the brain^{94,95}. The activity of most of these molecules is not known. Some of those that have been investigated show activities that have therapeutic potential; for example, arachidonoyl serine is a vasodilator⁹⁶ and is neuroprotective after brain injury as it reduces apoptosis⁹⁷. It leads to proliferation of neural progenitor cells *in vitro* and maintains these cells in an undifferentiated state *in vitro* and

in vivo. Although it does not bind to CB1 and CB2, its activity is blocked by CB2 antagonists⁹⁸. This raises questions, such as what is the relationship of such endocannabinoid-like compounds to the endocannabinoid system and what are the physiological roles of these molecules in the brain?

Pucci *et al.*⁹⁹ have investigated the possible epigenetic regulation of skin differentiation genes by phytocannabinoids⁹⁹. CBD was found to increase DNA methylation of the keratin 10 gene. Remarkably, CBD also reduced keratin 10 mRNA levels by a CB1-dependent mechanism. Thus, in this system, CBD is apparently a transcriptional repressor that can control cell proliferation and differentiation. As anandamide has also been found to have epigenetic properties¹⁰⁰, it is of interest to determine the extent, if any, of transcriptional control by endocannabinoids by epigenetic mechanisms.

Although various methods have been used to enhance endocannabinoid levels *in vivo* (even in patients)^{82,101}, neither anandamide nor 2-AG have been administered to humans. In addition, only a small number of clinical studies have been carried out using plant cannabinoids. A notable exception is the recent successful clinical trial using CBD in schizophrenic patients¹⁰¹. Although it is widely mentioned in the general media that cannabis with a high concentration of CBD is therapeutic in paediatric epilepsy and that 'medical marijuana' is indeed of value in such cases¹⁰², there have not been any recent clinical trials reported, although several such trials are ongoing (an anti-epileptic trial of CBD in adults was reported 34 years ago¹⁰³).

In a recent review, Pacher and Kunos⁸⁴ suggested that "modulating endocannabinoid system activity may have therapeutic potential in almost all diseases affecting humans". They supported this strong statement with a long list of examples, although these examples were mostly obtained *in vitro* or from *in vivo* experiments in animals⁸⁴. If this summary of effects is shown to reflect actions in human patients, is the endocannabinoid system going to bring a revolution in therapy? This might be the case as investigators are now able to target multiple cell-specific synthetic and biotransformation enzyme pathways that can adjust the levels of endocannabinoid ligands with some degree of tissue selectivity. In addition, aside from the agonist and antagonist ligands for cannabinoid receptors, researchers can now target cell type-specific allosteric modulators and receptor-associated proteins. Thus, there is great promise for the future of cannabinoid research.

Raphael Mechoulam is at the Institute for Drug Research, Medical Faculty, Hebrew University, Jerusalem, 91120, Israel.

Lumír O. Hanuš is at the Institute for Drug Research, Medical Faculty, Hebrew University, Jerusalem, 91120, Israel.

Roger Pertwee is at the Institute of Medical Sciences, University of Aberdeen, Aberdeen AB25 2ZD, Scotland, UK.

Allyn C. Howlett is at the Department of Physiology and Pharmacology, Wake Forest University Health Sciences, One Medical Center Blvd, Winston-Salem, North Carolina 27157, USA.

Correspondence to R.M.
e-mail: mechou@cc.huji.ac.il

doi:10.1038/nrn3811

Published online 15 October 2014

- Mechoulam, R. in *Cannabinoids as Therapeutic Agents*, (ed. Mechoulam, R.), 1–19 (CRC Press Inc., 1986).
- O'Shaughnessy, W. B. in *The Bengal Dispensary and Pharmacopoeia*, 579 (Bishop's College Press, 1841).
- Adams, R. *Marihuana. Harvey Lectures* **37**, 168–197 (1941–1942).
- Todd, A. R. *Hashish. Experientia* **2**, 55–60 (1946).
- Loewe, S. Cannabiswirkstoffe und Pharmacologie der Cannabinole. *Arch. Exp. Pathol. Pharmacol.* **211**, 175–193 (1950).
- Gaoni, Y. & Mechoulam, R. Isolation, structure and partial synthesis of an active constituent of hashish. *J. Amer. Chem. Soc.* **86**, 1646–1647 (1964).
- Mechoulam, R., McCallum, N. K. & Burstein, S. Recent advances in the chemistry and biochemistry of cannabis. *Chem. Rev.* **76**, 75–112 (1976).
- Agurell, S. *et al.* Pharmacokinetics and metabolism of Δ -1-tetrahydrocannabinol and other cannabinoids with emphasis on man. *Pharmacol. Rev.* **38**, 21–43 (1986).
- Pertwee, R. G. The ring test: a quantitative method for assessing the 'cataleptic' effect of cannabis in mice. *Br. J. Pharmacol.* **46**, 753–763 (1972).
- Pertwee, R. G. Cannabinoid pharmacology: the first 66 years. *Br. J. Pharmacol.* **147**, S163–S171 (2006).
- Pertwee, R. G. The central neuropharmacology of psychotropic cannabinoids. *Pharmacol. Ther.* **36**, 189–261 (1988).
- Pertwee, R. G. Emerging strategies for exploiting cannabinoid receptor agonists as medicines. *Br. J. Pharmacol.* **156**, 397–411 (2009).
- Mechoulam, R. *et al.* Stereochemical requirements for cannabinoid activity. *J. Med. Chem.* **23**, 1068–1072 (1980).
- Mechoulam, R. *et al.* Enantiomeric cannabinoids: stereospecificity of psychotropic activity. *Experientia* **44**, 762–764 (1988).
- Dewey, W. L. Cannabinoid pharmacology. *Pharmacol. Rev.* **38**, 151–178 (1986).
- Hollister, L. E. Health aspects of cannabis. *Pharmacol. Rev.* **38**, 1–20 (1986).
- Klee, W. A., Sharma, S. K. & Nirenberg, M. Opiate receptors as regulators of adenylate cyclase. *Life Sci.* **16**, 1869–1874 (1975).
- Nathanson, N. M., Klein, W. L. & Nirenberg, M. Regulation of adenylate cyclase activity mediated by muscarinic acetylcholine receptors. *Proc. Natl Acad. Sci. USA* **75**, 1788–1791 (1978).
- Sabol, S. L. & Nirenberg, M. Regulation of adenylate cyclase of neuroblastoma x glioma hybrid cells by α -adrenergic receptors. I. Inhibition of adenylate cyclase mediated by a receptors. *J. Biol. Chem.* **254**, 1913–1920 (1979).
- Howlett, A. C. & Fleming, R. M. Cannabinoid inhibition of adenylate cyclase. Pharmacology of the response in neuroblastoma cell membranes. *Mol. Pharmacol.* **26**, 532–538 (1984).
- Howlett, A. C. Inhibition of neuroblastoma adenylate cyclase by cannabinoid and nantadol compounds. *Life Sci.* **35**, 1803–1810 (1984).
- Howlett, A. C. Cannabinoid inhibition of adenylate cyclase. Biochemistry of the response in neuroblastoma cell membranes. *Mol. Pharmacol.* **27**, 429–436 (1985).

23. Howlett, A. C., Quail, J. M. & Khachatrian, L. L. Involvement of Gi in the inhibition of adenylate cyclase by cannabinimetic drugs. *Mol. Pharmacol.* **29**, 307–313 (1986).
24. Howlett, A. C., Champion, T. M., Wilken, G. H. & Mechoulam, R. Stereochemical effects of 11-OH- Δ^8 -tetrahydrocannabinol-dimethylheptyl to inhibit adenylate cyclase and bind to the cannabinoid receptor. *Neuropharmacology* **29**, 161–165 (1990).
25. Melvin, L. S. *et al.* Structure–activity relationships for cannabinoid receptor-binding and analgesic activity: studies of bicyclic cannabinoid analogs. *Mol. Pharmacol.* **44**, 1008–1015 (1993).
26. Melvin, L. S., Milne, G. M., Johnson, M. R., Wilken, G. H. & Howlett, A. C. Structure–activity relationships defining the ACD-tricyclic cannabinoids: cannabinoid receptor binding and analgesic activity. *Drug Des. Discov.* **13**, 155–166 (1995).
27. Devane, W. A., Dysarz F. A. 3rd, Johnson M. R., Melvin L. S. & Howlett A. C. Determination and characterization of a cannabinoid receptor in rat brain. *Mol. Pharmacol.* **34**, 605–613 (1988).
28. Gerard, C., Mollereau, C., Vassart, G. & Parmentier, M. Nucleotide sequence of a human cannabinoid receptor cDNA. *Nucleic Acids Res.* **18**, 7142 (1990).
29. Matsuda, L. A., Lolait, S. J., Brownstein, M. J., Young, A. C. & Bonner, T. I. Structure of a cannabinoid receptor and functional expression of the cloned cDNA. *Nature* **346**, 561–564 (1990).
30. Devane, W. A. *et al.* Isolation and structure of a brain constituent that binds to the cannabinoid receptor. *Science* **258**, 1946–1949 (1992).
31. Munro, S., Thomas, K. L. & Abu-Shaar, M. Molecular characterization of a peripheral receptor for cannabinoids. *Nature* **365**, 61–65 (1993).
32. Mechoulam, R. *et al.* Identification of an endogenous 2-monoglyceride, present in canine gut, that binds to cannabinoid receptors. *Biochem. Pharmacol.* **50**, 83–90 (1995).
33. Pertwee, R. G. *et al.* International Union of Basic and Clinical Pharmacology. LXXIX. Cannabinoid receptors and their ligands: beyond CB₁ and CB₂. *Pharmacol. Rev.* **62**, 588–631 (2010).
34. Pamplona, F. A. *et al.* Anti-inflammatory lipoxin A₂ is an endogenous allosteric enhancer of CB₁ cannabinoid receptor. *Proc. Natl Acad. Sci. USA* **109**, 21134–21139 (2012).
35. Bauer, M. *et al.* Identification and quantification of a new family of peptide endocannabinoids (pepcans) showing negative allosteric modulation at CB₁ receptors. *J. Biol. Chem.* **287**, 36944–36967 (2012).
36. Heimann, A. S. *et al.* Hemopressin is an inverse agonist of CB₁ cannabinoid receptors. *Proc. Natl Acad. Sci. USA* **104**, 20588–20593 (2007).
37. Huffman, J. W. *et al.* Synthesis and pharmacology of a very potent cannabinoid lacking a phenolic hydroxyl with high affinity for the CB₂ receptor. *J. Med. Chem.* **39**, 3875–3877 (1996).
38. Hanuš, L. *et al.* HU-308: A specific agonist for CB₂, a peripheral cannabinoid receptor. *Proc. Natl Acad. Sci. USA* **96**, 14228–14233 (1999).
39. Anand, P., Whiteside, G., Fowler, C. J. & Hohmann, A. G. Targeting CB₂ receptors and the endocannabinoid system for the treatment of pain. *Brain Res. Rev.* **60**, 255–266 (2009).
40. Fernandez-Ruiz, J., Pazos, M. R., Garcia-Arencibia, M., Sagredo, O. & Ramos, J. A. Role of CB₂ receptors in neuroprotective effects of cannabinoids. *Mol. Cell. Endocrinol.* **286**, S91–S96 (2008).
41. Marriott, K. S. & Huffman, J. W. Recent advances in the development of selective ligands for the cannabinoid CB₂ receptor. *Curr. Top. Med. Chem.* **8**, 187–204 (2008).
42. Pacher, P. & Mechoulam, R. Is lipid signaling through cannabinoid 2 receptors part of a protective system? *Progr. Lipid Res.* **50**, 193–211 (2011).
43. Horváth, B. *et al.* A new cannabinoid 2 receptor agonist HU-910 attenuates oxidative stress, inflammation, and cell death associated with hepatic ischemia/reperfusion injury. *Br. J. Pharmacol.* **165**, 2462–2478 (2012).
44. Di Marzo, V., De Petrocellis, L. & Bisogno, T. in *Cannabinoids. Handbook of Experimental Pharmacology* (ed. Pertwee, R. G.) **168**, 147–185 (Springer, 2005).
45. Wilson, R. I. & Nicoll R. A. Endogenous cannabinoids mediate retrograde signalling at hippocampal synapses. *Nature* **410**, 588–592 (2001).
46. Vaughan, C. W. & Christie, M. J. in *Cannabinoids. Handbook of Experimental Pharmacology* (ed. Pertwee, R. G.) **168**, 367–383 (Springer, 2005).
47. Ohno-Shosaku, T., Tanimura, A., Hashimoto, Y. & Kano, M. Endocannabinoids and retrograde modulation of synaptic transmission. *Neuroscientist* **18**, 119–132 (2012).
48. Alger, B. E. Endocannabinoids at the synapse a decade after the *dies mirabilis*: what we still do not know. *J. Physiol.* **590**, 2203–2212 (2012).
49. Gregg, L. C. *et al.* Activation of type 5 metabotropic glutamate receptors and diacylglycerol lipase- α initiates 2-arachidonoylglycerol formation and endocannabinoid-mediated analgesia. *J. Neurosci.* **32**, 9457–9468 (2012).
50. Rinaldi-Carmona, M. *et al.* SR141716A, a potent and selective antagonist of the brain cannabinoid receptor. *FEBS Lett.* **350**, 240–244 (1994).
51. Rinaldi-Carmona, M. *et al.* Characterization and distribution of binding sites for [3H]-SR141716A, a selective brain (CB₁) cannabinoid receptor antagonist, in rodent brain. *Life Sci.* **58**, 1239–1247 (1996).
52. Mathews, W. B. *et al.* Biodistribution of [¹⁸F] SR144385 and [¹⁸F] SR147963: selective radioligands for positron emission tomographic studies of brain cannabinoid receptors. *Nucl. Med. Biol.* **27**, 757–762 (2000).
53. Bouaboula, M. *et al.* Stimulation of cannabinoid receptor CB₁ induces krox-24 expression in human astrocytoma cells. *J. Biol. Chem.* **270**, 13973–13980 (1995).
54. Bouaboula, M. *et al.* Activation of mitogen-activated protein kinases by stimulation of the central cannabinoid receptor CB₁. *Biochem. J.* **312**, 637–641 (1995).
55. Bouaboula, M. *et al.* A selective inverse agonist for central cannabinoid receptor inhibits mitogen-activated protein kinase activation stimulated by insulin or insulin-like growth factor 1. Evidence for a new model of receptor/ligand interactions. *J. Biol. Chem.* **272**, 22330–22339 (1997).
56. Compton, D. R., Aceto, M. D., Lowe, J. & Martin, B. R. *In vivo* characterization of a specific cannabinoid receptor antagonist (SR141716A): inhibition of Δ^9 -tetrahydrocannabinol-induced responses and apparent agonist activity. *J. Pharmacol. Exp. Ther.* **277**, 586–594 (1996).
57. Gueudet, C., Santucci, V., Rinaldi-Carmona, M., Soubrie, P. & Le Fur, G. The CB₁ cannabinoid receptor antagonist SR141716A affects A9 dopamine neuronal activity in the rat. *Neuroreport* **6**, 1421–1425 (1995).
58. Perio, A. *et al.* Central mediation of the cannabinoid cue: activity of a selective CB₁ antagonist, SR141716A. *Behav. Pharmacol.* **7**, 65–71 (1996).
59. Aceto, M. D., Scates, S. M., Lowe, J. A. & Martin, B. R. Cannabinoid precipitated withdrawal by the selective cannabinoid receptor antagonist, SR141716A. *Eur. J. Pharmacol.* **282**, R1–R2 (1995).
60. Tsou, K., Patrick, S. L. & Walker, J. M. Physical withdrawal in rats tolerant to Δ^9 -tetrahydrocannabinol precipitated by a cannabinoid receptor antagonist. *Eur. J. Pharmacol.* **280**, R13–R15 (1995).
61. Felder, C. C. *et al.* LY320135, a novel cannabinoid CB₁ receptor antagonist, unmasks coupling of the CB₁ receptor to stimulation of cAMP accumulation. *J. Pharmacol. Exp. Ther.* **284**, 291–297 (1998).
62. Meschler, J. P., Kraichely, D. M., Wilken, G. H. & Howlett, A. C. Inverse agonist properties of N-(piperidin-1-yl)-5-(4-chlorophenyl)-1-(2,4-dichlorophenyl)-4-methyl-1H-pyrazole-3-carboxamide HCl (SR141716A) and 1-(2-chlorophenyl)-4-cyano-5-(4-methoxyphenyl)-1H-pyrazole-3-carboxylic acid phenylamide (CP-272871) for the CB₁ cannabinoid receptor. *Biochem. Pharmacol.* **60**, 1315–1323 (2000).
63. Cosenza, M. *et al.* Locomotor activity and occupancy of brain cannabinoid CB₁ receptors by the antagonist/inverse agonist AM281. *Synapse* **38**, 477–482 (2000).
64. Lan, R. *et al.* Design and synthesis of the CB₁ selective cannabinoid antagonist AM281: a potential human SPECT ligand. *AAPS Pharm. Sci.* **1**, E4 (1999).
65. Pertwee, R. *et al.* AM630, a competitive cannabinoid receptor antagonist. *Life Sci.* **56**, 1949–1955 (1995).
66. Rinaldi-Carmona, M. *et al.* SR144528, the first potent and selective antagonist of the CB₂ cannabinoid receptor. *J. Pharmacol. Exp. Ther.* **284**, 644–650 (1998).
67. Cascio, M. G. *et al.* *In vitro* and *in vivo* pharmacological characterization of two novel selective cannabinoid CB₂ receptor inverse agonists. *Pharmacol. Res.* **61**, 349–354 (2010).
68. Miller, A. M. & Stella, N. CB₂ receptor-mediated migration of immune cells: it can go either way. *Br. J. Pharmacol.* **153**, 299–308 (2008).
69. Fernandez-Ruiz, J. *et al.* Cannabinoid CB₂ receptor: a new target for controlling neural cell survival? *Trends Pharmacol. Sci.* **28**, 39–45 (2007).
70. Wright, K. L., Duncan, M. & Sharkey, K. A. Cannabinoid CB₂ receptors in the gastrointestinal tract: a regulatory system in states of inflammation. *Br. J. Pharmacol.* **153**, 263–270 (2008).
71. Lunn, C. A. *et al.* Biology and therapeutic potential of cannabinoid CB₂ receptor inverse agonists. *Br. J. Pharmacol.* **153**, 226–239 (2008).
72. Scheen, A. J. *et al.* Efficacy and tolerability of rimonabant in overweight or obese patients with type 2 diabetes: a randomised controlled study. *Lancet* **368**, 1660–1672 (2006).
73. Nissen, S. E. *et al.* Effect of rimonabant on progression of atherosclerosis in patients with abdominal obesity and coronary artery disease: the STRADIVARIUS randomized controlled trial. *JAMA* **299**, 1547–1560 (2008).
74. Van Gaal, L. F., Rissanen, A. M., Scheen, A. J., Ziegler, O. & Rossner, S. Effects of the cannabinoid-1 receptor blocker rimonabant on weight reduction and cardiovascular risk factors in overweight patients: 1-year experience from the RIO-Europe study. *Lancet* **365**, 1389–1397 (2005).
75. Moreira, F. A., Grieb, M. & Lutz, B. Central side-effects of therapies based on CB₁ cannabinoid receptor agonists and antagonists: focus on anxiety and depression. *Best Pract. Res. Clin. Endocrinol. Metab.* **23**, 133–144 (2009).
76. Nathan, P. J., O'Neill, B. V., Napolitano, A. & Bullmore, E. T. Neuropsychiatric adverse effects of centrally acting antipsychotic drugs. *CNS Neurosci. Ther.* **17**, 490–505 (2011).
77. Di Marzo, V. & Despres, J. P. CB₁ antagonists for obesity — what lessons have we learned from rimonabant? *Nature Rev. Endocrinol.* **5**, 633–638 (2009).
78. Kirilly, E., Gonda, X. & Bagdy, G. CB₁ receptor antagonists: new discoveries leading to new perspectives. *Acta Physiol.* **205**, 41–60 (2012).
79. Tam, J. *et al.* Peripheral cannabinoid-1 receptor inverse agonism reduces obesity by reversing leptin resistance. *Cell. Metab.* **16**, 167–179 (2012).
80. Lazary, J., Juhasz, G., Hunyadi, L. & Bagdy, G. Personalized medicine can pave the way for the safe use of CB₁ receptor antagonists. *Trends Pharmacol. Sci.* **32**, 270–280 (2011).
81. Blankman, J. L. & Cravatt, B. F. Chemical probes of endocannabinoid metabolism. *Pharmacol. Rev.* **65**, 849–871 (2013).
82. Pertwee, R. G. Elevating endocannabinoid levels: pharmacological strategies and potential therapeutic applications. *Proc. Nutr. Soc.* **73**, 96–105 (2014).
83. Pertwee, R. G. The therapeutic potential of drugs that target cannabinoid receptors or modulate the tissue levels or actions of endocannabinoids. *AAPS J.* **7**, E625–E654 (2005).
84. Pacher, P. & Kunos, G. Modulating the endocannabinoid system in human health and disease — successes and failures. *FEBS J.* **280**, 1918–1943 (2013).
85. Pertwee, R. G. The diverse CB₁ and CB₂ receptor pharmacology of three plant cannabinoids: Δ^9 -tetrahydrocannabinol, cannabidiol and Δ^9 -tetrahydrocannabinol. *Br. J. Pharmacol.* **153**, 199–215 (2008).
86. Pertwee, R. G. Receptors and channels targeted by synthetic cannabinoid receptor agonists and antagonists. *Curr. Med. Chem.* **17**, 1360–1381 (2010).
87. Pertwee, R. G. & Cascio, M. G. in *Handbook of Cannabis* (ed. Pertwee, R. G.) 115–136 (Oxford University Press, 2014).
88. Pertwee, R. G. Pharmacology of cannabinoid receptor ligands. *Curr. Med. Chem.* **6**, 635–664 (1999).
89. McHugh, D., Page, J., Dunn, E. & Bradshaw, H. B. Δ^9 -Tetrahydrocannabinol and N-arachidonyl glycine are full agonists at GPR18 receptors and induce migration in human endometrial HEC-1B cells. *Br. J. Pharmacol.* **165**, 2414–2424 (2012).
90. De Petrocellis, L. *et al.* Effects of cannabinoids and cannabinoid-enriched *Cannabis* extracts on TRP channels and endocannabinoid metabolic enzymes. *Br. J. Pharmacol.* **163**, 1479–1494 (2011).
91. Di Marzo, V. A brief history of cannabinoid and endocannabinoid pharmacology as inspired by the work of British scientists. *Trends Pharmacol. Sci.* **27**, 134–140 (2006).

92. Howlett, A. C., Blume, L. C. & Dalton, G. D. CB1 cannabinoid receptors and their associated proteins. *Curr. Med. Chem.* **17**, 1382–1393 (2010).
93. Smith, T. H., Sim-Selley, L. J. & Selley, D. E. Cannabinoid CB1 receptor-interacting proteins: novel targets for central nervous system drug discovery? *Br. J. Pharmacol.* **160**, 454–466 (2010).
94. Tan, B. *et al.* Targeted lipidomics: discovery of new fatty acyl amides. *AAPS J.* **8**, E461–E465 (2006).
95. Tan, B. *et al.* Identification of endogenous acyl amino acids based on a targeted lipidomics approach. *J. Lipid Res.* **51**, 112–119 (2010).
96. Milman, G. *et al.* N-Arachidonoyl L-serine, a novel endocannabinoid-like brain constituent with vasodilatory properties. *Proc. Natl Acad. Sci. USA* **103**, 2428–2433 (2006).
97. Cohen-Yeshurun, A. *et al.* N-Arachidonoyl-L-serine is neuroprotective after traumatic brain injury by reducing apoptosis. *J. Cereb. Blood Flow Metab.* **31**, 1768–1777 (2011).
98. Cohen-Yeshurun, A. *et al.* N-arachidonoyl-L-serine (AraS) possesses pro-neurogenic properties *in vitro* and *in vivo* following traumatic brain injury. *J. Cereb. Blood Flow Metab.* **33**, 1242–1250 (2013).
99. Pucci, M. *et al.* Epigenetic control of skin differentiation genes by phytocannabinoids. *Br. J. Pharmacol.* **170**, 581–591 (2013).
100. Pasquariello, N., Oddi, S., Malaponti, M. & Maccarrone, M. Regulation of gene transcription and keratinocyte differentiation by anandamide. *Vitam. Horm.* **81**, 441–467 (2009).
101. Leweke, F. M. *et al.* Cannabidiol enhances anandamide signaling and alleviates psychotic symptoms of schizophrenia. *Transl. Psychiatry* **2**, e94 (2012).
102. Porter, B. E. & Jacobson, C. Report of a parent survey of cannabidiol-enriched cannabis use in pediatric treatment-resistant epilepsy. *Epilepsy Behav.* **29**, 574–577 (2013).
103. Cunha, J. M. *et al.* Chronic administration of cannabidiol to healthy volunteers and epileptic patients. *Pharmacol.* **21**, 175–185 (1980).
104. Wood, T. B., Spivey, W. T. N. & Easterfield, T. H. Cannabinol. Part I. *J. Chem. Soc.* **75**, 20–36 (1899).
105. Cahn, R. S. Cannabis indica resin, Part, III The constitution of Cannabinol. *J. Chem. Soc.* 1342–1353 (1932).
106. Mechoulam, R. & Shvo, Y. The structure of cannabidiol. *Tetrahedron* **19**, 2073–2078 (1963).
107. Sugiura, T. *et al.* 2-Arachidonoylglycerol: a possible endogenous cannabinoid receptor ligand in brain. *Biochem. Biophys. Res. Commun.* **215**, 89–97 (1995).
108. Cravatt, B. F. *et al.* Molecular characterization of an enzyme that degrades neuromodulatory fatty-acid amides. *Nature* **384**, 83–87 (1996).
109. Zygmunt, P. M. *et al.* Vanilloid receptors on sensory nerves mediate the vasodilator action of anandamide. *Nature* **400**, 452–457 (1999).
110. Smart, D. *et al.* The endogenous lipid anandamide is a full agonist at the human vanilloid receptor (hVR1). *Br. J. Pharmacol.* **129**, 227–230 (2000).
111. Piomelli, D. A fatty gut feeling. *Trends Endocrinol. Metab.* **24**, 332–341 (2013).
112. Bandiera, T., Ponzano, S. & Piomelli, D. Advances in the discovery of N-acyl ethanolamine acid amidase inhibitors. *Pharmacol. Res.* **86C**, 11–17 (2014).
113. Schlosburg, J. E. *et al.* Prolonged monoacylglycerol lipase blockade causes equivalent cannabinoid receptor type 1 receptor-mediated adaptations in fatty acid amide hydrolase wild-type and knockout mice. *J. Pharmacol. Exp. Ther.* **350**, 196–204 (2014).
114. Galve-Roperh, I. *et al.* Cannabinoid receptor signaling in progenitor/stem cell proliferation and differentiation. *Prog. Lipid Res.* **52**, 633–650 (2013).
115. Katona, I. & Freund, T. F. Multiple functions of endocannabinoid signaling in the brain. *Annu. Rev. Neurosci.* **35**, 529–558 (2012).
116. Piomelli, D. & Sasso, O. Peripheral gating of pain signals by endogenous lipid mediators. *Nature Neurosci.* **17**, 164–174 (2014).
117. Syed, Y. Y., McKeage, K. & Scott, L. J. Δ^9 -tetrahydrocannabinol/cannabidiol (Sativex): a review of its use in patients with moderate to severe spasticity due to multiple sclerosis. *Drugs* **74**, 563–578 (2014).
118. Roitman, P., Mechoulam, R., Cooper-Kazaz, R. & Shalev, A. Preliminary, open-label, pilot study of add-on oral $\delta(9)$ -tetrahydrocannabinol in chronic post-traumatic stress disorder. *Clin. Drug Investig.* **34**, 587–591 (2014).
119. Bisogno, T. *et al.* Cloning of the first sn1-DAG lipases points to the spatial and temporal regulation of endocannabinoid signaling in the brain. *J. Cell Biol.* **163**, 463–468 (2003).

Acknowledgements

Research in the laboratory of R.M. was supported by the Kessler Family Foundation, Boston, USA, and by a grant from US National Institute on Drug Abuse (NIDA), DA-9789. The research of R.P. was supported by NIDA grants DA-3934, DA-9789 and DA-3672 and GW Pharmaceuticals and the research of A.H. was supported by NIDA grant DA-3690.

Competing interests statement

The authors declare no competing interests.

ORIGINAL ARTICLE

Genetic predisposition to schizophrenia associated with increased use of cannabis

RA Power^{1,2}, KJH Verweij³, M Zuhair¹, GW Montgomery⁴, AK Henders⁴, AC Heath⁵, PAF Madden⁵, SE Medland⁴, NR Wray² and NG Martin⁴

Cannabis is the most commonly used illicit drug worldwide. With debate surrounding the legalization and control of use, investigating its health risks has become a pressing area of research. One established association is that between cannabis use and schizophrenia, a debilitating psychiatric disorder affecting ~1% of the population over their lifetime. Although considerable evidence implicates cannabis use as a component cause of schizophrenia, it remains unclear whether this is entirely due to cannabis directly raising risk of psychosis, or whether the same genes that increases psychosis risk may also increase risk of cannabis use. In a sample of 2082 healthy individuals, we show an association between an individual's burden of schizophrenia risk alleles and use of cannabis. This was significant both for comparing those who have ever versus never used cannabis ($P = 2.6 \times 10^{-4}$), and for quantity of use within users ($P = 3.0 \times 10^{-3}$). Although directly predicting only a small amount of the variance in cannabis use, these findings suggest that part of the association between schizophrenia and cannabis is due to a shared genetic aetiology. This form of gene–environment correlation is an important consideration when calculating the impact of environmental risk factors, including cannabis use.

Molecular Psychiatry (2014) **19**, 1201–1204; doi:10.1038/mp.2014.51; published online 24 June 2014

INTRODUCTION

During the last quarter of the 20th century, cannabis use has increased to become the most widely used illicit drug in the world.¹ It is well established that cannabis use is much higher among schizophrenic patients than in the general population.² Cannabis intoxication can lead to an acute transient psychotic episode and produce short-term exacerbations of pre-existing psychotic symptoms,^{3–5} an association that has been confirmed through the experimental administration of tetrahydrocannabinol.^{6,7} Meta-analyses of prospective studies have found that cannabis use increases the likelihood of developing a psychotic illness by a factor of roughly two.^{8–11} A dose response effect has been demonstrated,^{12–14} and use in adolescence has been associated with the greatest risk.¹⁵ Given the large health burden from schizophrenia and other psychotic disorders,¹⁶ the view that cannabis use is a component cause of schizophrenia has heavily influenced discussion over the legislation surrounding cannabis use.

However, the relationship between schizophrenia and cannabis use may be more complicated than it initially seems. Despite a clear association between the two, the possibility of reverse causation has not been entirely excluded. Some small studies have suggested that it is in fact psychosis that is a risk factor for cannabis use, as those on a psychotic spectrum are more likely to experiment with drugs.^{17,18} The strongest evidence comes from Ferdinand *et al.*¹⁹ who found that the association was bidirectional, as cannabis-naïve children with prodromal psychotic episodes had greater incidence of later cannabis use. However, a similarly sized study failed to replicate this finding.²⁰ There is also the possibility of attempts by patients at self-medication, as it

has been suggested that cannabis use can reduce negative and affective symptoms in patients with an established psychotic disorder.^{21–23}

The issue is further complicated by tentative evidence for interactions between cannabis use and genetic risk variants for schizophrenia.²⁴ Schizophrenia is known to be highly heritable with up to 80% of the variance explained by additive genetic effects,²⁵ and as sample sizes have increased a growing number of genetic risk variants have been identified.^{26,27} Interactions between risk variants and cannabis use might explain why some individuals experience psychosis while others do not. However, cannabis use itself has been reported to be heritable,^{28–30} although no genetic risk variants have been identified.³¹ It is unclear to what extent the heritability of cannabis use results from shared heritability with other behavioural phenotypes such as schizophrenia predicting its use.

Here we test for such genetic overlap directly, and aim to discern the direction of causation between cannabis use and schizophrenia. Within a sample of 2082 healthy individuals, we tested to see whether polygenic risk scores for schizophrenia predict cannabis use. Polygenic risk scores reflect the cumulative burden of risk alleles carried by an individual as identified in a previous genome-wide association study (GWAS),³² here of 13 833 schizophrenia cases and 18 310 controls.²⁷ Such an association with cannabis use would suggest that those genetically predisposed to schizophrenia use cannabis more frequently. This would mean that the association between schizophrenia and cannabis use is not simply one of an environmental risk factor, but rather involves gene–environment correlation, as individuals

¹MRC Social, Genetic & Developmental Psychiatry Centre, Institute of Psychiatry, King's College London, London, UK; ²Queensland Brain Institute, The University of Queensland, St Lucia, QLD, Australia; ³Department of Developmental Psychology and EMGO Institute for Health and Care Research, VU University, Amsterdam, The Netherlands; ⁴QIMR Berghofer Medical Research Institute, Brisbane, QLD, Australia and ⁵Department of Psychiatry, Washington University School of Medicine, St Louis, MO, USA. Correspondence: R Power, MRC Social, Genetic & Developmental Psychiatry Centre, Institute of Psychiatry, King's College London, DeCrespigny Park, Denmark Hill, London SE5 8AF, UK. E-mail: robert.r.power@kcl.ac.uk
Received 22 January 2014; revised 18 March 2014; accepted 22 April 2014; published online 24 June 2014

choose and shape their own environment based on their own innate preferences.

MATERIALS AND METHODS

The data used in this study come from the Australian Twin Registry. Data were obtained from two studies in which twins and their families participated in semi-structured diagnostic telephone interviews aimed primarily at assessing psychiatric health. Informed consent was obtained from all participants.

Sample 1 consisted of 6265 individuals aged between 23 and 39 years (mean = 29.9 ± 2.5) interviewed between 1996 and 2000. Participants were members of the young adult cohort, a volunteer panel of twins born between 1964 and 1971. The interview was based on a modified version of the SSAGA (Semi-Structured Assessment of the Genetics of Alcoholism³³). Detailed information about the sample recruitment, the study procedure and the measures can be found elsewhere.³⁴ Sample 2 comprised 9688 individuals aged between 18 and 91 years (mean = 46.3 ± 11.3) interviewed between 2001 and 2005. Participants were members of the older and younger adult cohort of Australian twin pairs (born between 1895 and 1964, and between 1964 and 1971, respectively). A subset of this sample was ascertained based on large sibship size, or having a relative with nicotine or alcohol dependence. The interview used for this sample was also based on a modified version of the SSAGA. Further details about the sample and assessment can be found in Heath et al.³⁵

A subset of the participants ($N=1866$; 11.7%) participated in both studies, in which case we used data from the last assessment. The combined phenotypic sample consisted of 14 087 individuals, of whom 7172 were genotyped. In both studies, twins were asked the same items about cannabis use: (1) did you ever use marijuana?, (2) how old were you the very first time you tried marijuana (not counting the times you took it as prescribed)? and (3) how many times in your life have you used marijuana (do not count times when you used a drug prescribed for you and took the prescribed dose). Ever use was measured on a dichotomous scale (ever versus never), whereas age at initiation and quantity of use were open questions. Table 1 shows the prevalence of cannabis use for individuals included in the present study.

Genotype data were obtained using three different Illumina single nucleotide polymorphism (SNP) genotyping platforms (317K, HumanCNV370-Quadv3, Human CNV370v1 and Human610-Quad). Standard quality control procedures were applied as outlined previously,³⁶ including checks for ancestry outliers, Hardy-Weinberg equilibrium ($P < 10^{-6}$), Mendelian errors, call rate, genotypic missingness ($>5\%$), individual missingness ($>5\%$) and minor allele frequency (<0.01). Individuals were pruned on relatedness, removing one individual from each pair with relatedness >0.05 , as determined from genetic data. The final sample therefore comprised 2082 'unrelated' individuals (see Table 1 for sample details).

Polygenic risk scores were constructed using the P -values and \log_{10} odds ratios from the most recent large GWAS of schizophrenia, a meta-analysis of the Psychiatric Genomics Consortium's studies with additional Swedish samples totalling 13 833 cases and 18 310 controls.²⁷ SNPs were pruned for linkage disequilibrium using P -value informed clumping in PLINK,³⁷ using a cutoff of $R^2=0.25$ within 200 kb window. The major histocompatibility complex region of the genome was excluded, due to its complex linkage disequilibrium structure. After linkage disequilibrium pruning, 147 830 SNPs remained. Multiple scores were generated for each individual using the PLINK score option and based on top SNPs from the schizophrenia GWAS using varying significance thresholds ($P=0.0001$, 0.001, 0.01, 0.05, 0.1, 0.2, 0.3, 0.4, 0.5 and 1.0). Polygenic risk scores were tested for association with a binary ever versus never used cannabis and two quantitative traits for quantity of use and age at first use, in logistic and linear regressions, respectively. These analyses were corrected for the

first 10 ancestry-informative principal components, genotyping platform, sex, age, age squared and sex by age. Analysis was performed in STATA.³⁸

RESULTS

After pruning, 2082 unrelated individuals remained in our sample with both genotype and phenotype measures. Within the sample, 1011 individuals (48.6%) had ever used cannabis, of whom 997 had data on quantity of use. Mean number of usages of cannabis over lifetime was 62.7 (95% CI 53.8–71.6), and the mean age of initiation of use was 20.1 (95% CI 19.7–20.5). Males showed higher rates of use than females, 53.5% compared with 43.9% ($P < 0.001$), although no significant difference in age at initiation. Table 1 shows the summary statistics for the sample.

Polygenic risk scores for schizophrenia showed positive associations for ever versus never use of cannabis across all P -value thresholds, with the strongest association for those SNPs with P -values of 0.01 or below in the original schizophrenia GWAS (see Figure 1, $R^2=0.47\%$, $P=2.6 \times 10^{-4}$). Significant associations were also seen in the analysis of quantity of cannabis use for 9 of the 10 SNP cutoffs, with the top association seen for those SNPs with $P \leq 0.05$ for schizophrenia ($R^2=0.85\%$, $P=0.003$). No association was seen with age at initiation of use, although the association with quantity of use remained significant when number of years of usage was accounted for (results not shown).

As a secondary analysis, polygenic risk score for schizophrenia risk alleles with $P \leq 0.01$ (the threshold with the greatest association in the primary analysis) was examined within 990 twin pairs (608 dizygotic and 382 monozygotic) where data on cannabis use of both twins was available. Taking the mean polygenic risk score within each twin pair, an ordinal regression was performed to predict whether neither ($n=272$), one ($n=273$) or both twins ($n=445$) were cannabis users. After correcting for age, sex and zygosity, a significant association was observed ($P=0.001$). Those twin pairs where both reported using cannabis had the greatest burden of schizophrenia risk alleles, pairs with only one user were found to have an intermediate level and the

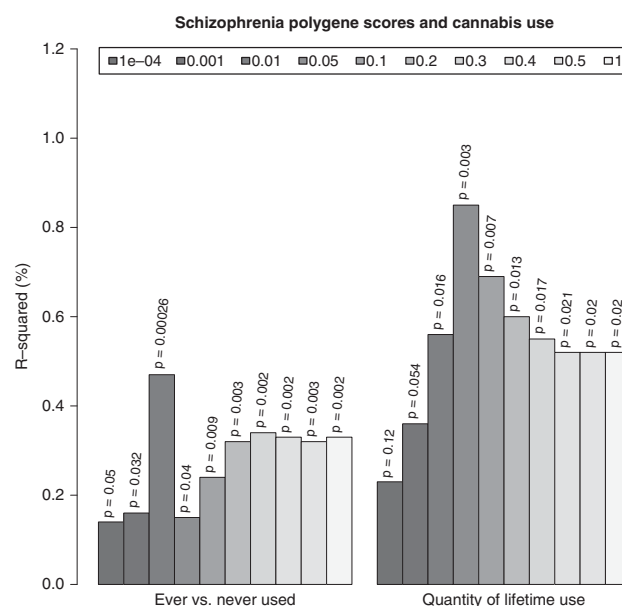


Figure 1. Results of polygenic risk scores for schizophrenia predicting variance explained (R^2) in cannabis use as both a binary trait of ever versus never, and as a quantitative trait of lifetime use within only users. Polygenic scores were created using different cutoffs for the inclusion of risk variants for schizophrenia, ranging from $P=0.0001$ to 1.0.

Table 1. Summary statistics of sample for cannabis use traits

	Users	Non-users
<i>N</i>	1011	1071
Mean age (s.e.)	41.3 (0.23)	53.0 (0.37)
Percentage female (%)	46.5	56.0
Mean age at initiation (s.e.)	19.6 (0.06)	—
Mean number of uses over lifetime (s.e.)	62.7 (4.56)	—

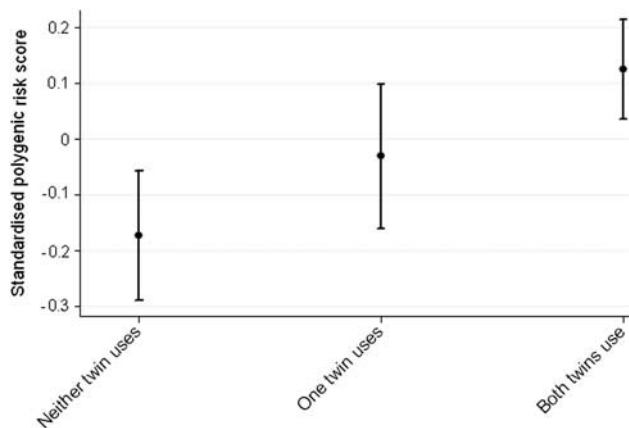


Figure 2. Mean standardized schizophrenia polygenic risk scores for pairs of twins when neither ($n=272$), one ($n=273$) or both twins ($n=445$) had reported use of cannabis. An ordinal regression reported a significant association ($P=0.001$).

lowest burden was found in pairs where neither twin reported use (see Figure 2).

DISCUSSION

Our results show that to some extent the association between cannabis and schizophrenia is due to a shared genetic aetiology across common variants. They suggest that individuals with an increased genetic predisposition to schizophrenia are both more likely to use cannabis and to use it in greater quantities. This is not to say that there is no causal relationship between use of cannabis and risk of schizophrenia, but it does establish that at least part of the association may be due to causal relationship in the opposite direction. Although the variance in cannabis use explained by schizophrenia polygenic risk scores is small, it is in line with other cross-phenotype analyses, largely due to the polygenic risk scores for schizophrenia predicting only ~7% of the variation for schizophrenia itself. Previous associations between polygenic risk scores for schizophrenia and other psychiatric illnesses, such as bipolar disorder, major depression and autism,³⁹ have shown effects of similar sizes. Further research will be needed to see whether the genetic overlap observed here is specific to cannabis use or is present across illicit drug use and addiction phenotypes, data for which was not widely available in this sample. For now, these findings have important implications for the current perception of cannabis use as a risk factor for schizophrenia, and other psychotic disorders.

However, it is worth noting that this association, if true, does not rule out the possibility of cannabis independently being a risk factor for schizophrenia. A bidirectional association between cannabis use and psychosis has previously been suggested.⁴⁰ Further, one caveat to interpreting the direction of causation concerns the discovery sample used to identify schizophrenia risk alleles. The schizophrenia GWAS sample will likely include many more cannabis users among cases than controls. This may lead to an excess of causal SNPs associated with cannabis use, as opposed to schizophrenia itself, identified as schizophrenia risk alleles. Only if the discovery schizophrenia sample was comprised entirely of non-cannabis users could causation be inferred without any risk of confounding. This is an important consideration as to whether polygenic risk scores overestimate individuals' un-modifiable genetic risk by including their genetic predisposition to modifiable environmental risk factors.

These results highlight the blurring between behavioural phenotypes and environment, and have wider implications for how we perceive supposedly environmental risks for disease.

Individuals select their own environments based on their innate and learned preferences, and have their environments react to their own behaviour. Further, parents pass down both genes and environment to their children. All of these can contribute to gene–environment correlation, particularly with respect to behavioural traits. Several studies have shown that supposedly environmental risk factors such as urbanicity, religiosity and stressful life events have heritable components to them.^{41–43} The existence of heritability for supposedly environmental risk factors does not mean they are inevitable, only that causality is more complicated to discern. Future studies will need to explore the matching of cases and controls on environmental risk variants to fully disentangle causation. This can be supplemented exploring the generation of polygenic risk scores for environmental risk factors, and their role in predicting disease status. The wider availability of genetic data in richly phenotyped samples should allow for the integration of genetics into an epidemiological framework, and so the discovery of gene–environment correlations where they exist.

With ongoing debate over the legalization of cannabis and the potential health risks it poses, understanding the association between its use and schizophrenia is a priority. It has previously been suggested that, even assuming an entirely causal relationship, the required reduction in the number of cannabis users to prevent one case of schizophrenia is in the thousands.⁴⁴ Our findings here highlight the possibility that this association might be bidirectional in causation, and that the risks of cannabis use could be overestimated. This is an important subtlety to consider when calculating the economic and health impact of cannabis use.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ACKNOWLEDGMENTS

Robert Power was funded by the Medical Research Council and the National Institute for Health Research (NIHR) Biomedical Research Centre at South London and Maudsley NHS Foundation Trust and King's College London. The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health. This work was supported by National Institutes of Health Grants AA07535, AA07580, AA07728, AA10249, AA13320, AA13321, AA14041, AA11998, AA17688, DA012854, DA018267, DA018660, DA23668 and DA019951; by Grants from the Australian National Health and Medical Research Council (241944, 339462, 389927, 389875, 389891, 389892, 389938, 442915, 442981, 496739, 552485, 552498, 6136022, 628911 and 1047956); by Grants from the Australian Research Council (A7960034, A79906588, A79801419, DP0770096, DP0212016 and DP0343921); and by the 5th Framework Programme (FP-5) GenomeUtwinn Project (QLG2-CT-2002-01254). This research was further supported by the Centre for Research Excellence on Suicide Prevention (CRESP—Australia). KJHV is supported by the Netherlands Organization for Health Research and Development, ZonMW 31160212. We thank Richard Parker, Soad Hancock, Judith Moir, Sally Rodda, Pieta-Maree Shertock, Heather Park, Jill Wood, Pam Barton, Fran Husband, Adele Somerville, Ann Eldridge, Marlene Grace, Kerrie McAloney, Lisa Bowdler, Alexandre Todorov, Steven Crooks, David Smyth, Harry Beeby and Daniel Park. Finally, we thank the twins and their families for their participation.

REFERENCES

- 1 United Nations Office on Drugs and Crime. *World Drug Report 2013*. United Nations: New York, NY, USA, 2013.
- 2 Green B, Young R, Kavanagh D. Cannabis use and misuse prevalence among people with psychosis. *Br J Psychiatry* 2005; **187**: 306–313.
- 3 McGuire PK, Jones P, Harvey I, Bebbington P, Toone B, Lewis S *et al*. Cannabis and acute psychosis. *Schizophr Res* 1994; **13**: 161–167.
- 4 Thornicroft G. Cannabis and psychosis. Is there epidemiological evidence for an association?. *Br J Psychiatry* 1990; **157**: 25–33.
- 5 Tien AY, Anthony JC. Epidemiological analysis of alcohol and drug use as risk factors for psychotic experiences. *J Nerv Ment Dis* 1990; **178**: 473–480.
- 6 D'Souza DC, Perry E, MacDougall L, Ammerman Y, Cooper T, Wu YT *et al*. The psychotomimetic effects of intravenous delta-9-tetrahydrocannabinol in healthy

- individuals: implications for psychosis. *Neuropsychopharmacology* 2004; **29**: 1558–1572.
- 7 Morrison PD, Zois V, McKeown DA, Lee TD, Holt DW, Powell JF et al. The acute effects of synthetic intravenous Delta9-tetrahydrocannabinol on psychosis, mood and cognitive functioning. *Psychol Med* 2009; **39**: 1607–1616.
 - 8 van Os J, Bak M, Hanssen M, Bijl RV, de Graaf R, Verdoux H. Cannabis use and psychosis: a longitudinal population-based study. *Am J Epidemiol* 2002; **156**: 319–327.
 - 9 Regier DA, Farmer ME, Rae DS, Locke BZ, Keith SJ, Judd LL et al. Comorbidity of mental disorders with alcohol and other drug abuse. Results from the Epidemiologic Catchment Area (ECA) Study. *JAMA* 1990; **264**: 2511–2518.
 - 10 Arseneault L, Cannon M, Witton J, Murray RM. Causal association between cannabis and psychosis: examination of the evidence. *Br J Psychiatry* 2004; **184**: 110–117.
 - 11 McGrath J, Welham J, Scott J, Varghese D, Degenhardt L, Hayatbakhsh MR et al. Association between cannabis use and psychosis-related outcomes using sibling pair analysis in a cohort of young adults. *Arch Gen Psychiatry* 2010; **67**: 440–447.
 - 12 Zammit S, Allebeck P, Andreasson S, Lundberg I, Lewis G. Self reported cannabis use as a risk factor for schizophrenia in Swedish conscripts of 1969: historical cohort study. *BMJ* 2002; **325**: 1199.
 - 13 Henquet C, Krabbendam L, Spauwen J, Kaplan C, Lieb R, Wittchen HU et al. Prospective cohort study of cannabis use, predisposition for psychosis, and psychotic symptoms in young people. *BMJ* 2005; **330**: 11.
 - 14 Moore TH, Zammit S, Lingford-Hughes A, Barnes TR, Jones PB, Burke M et al. Cannabis use and risk of psychotic or affective mental health outcomes: a systematic review. *Lancet* 2007; **370**: 319–328.
 - 15 Arseneault L, Cannon M, Poulton R, Murray R, Caspi A, Moffitt TE. Cannabis use in adolescence and risk for adult psychosis: longitudinal prospective study. *BMJ* 2002; **325**: 1212–1213.
 - 16 Vos T, Flaxman AD, Naghavi M, Lozano R, Michaud C, Ezzati M et al. Years lived with disability (YLDs) for 1160 sequelae of 289 diseases and injuries 1990–2010: a systematic analysis for the Global Burden of Disease Study 2010. *Lancet* 2012; **380**: 2163–2196.
 - 17 Addington J, Duchak V. Reasons for substance use in schizophrenia. *Acta Psychiatr Scand* 1997; **96**: 329–333.
 - 18 Spencer C, Castle D, Michie PT. Motivations that maintain substance use among individuals with psychotic disorders. *Schizophr Bull* 2002; **28**: 233–247.
 - 19 Ferdinand RF, Sondeijker F, van der Ende J, Selten JP, Huizink A, Verhulst FC. Cannabis use predicts future psychotic symptoms, and vice versa. *Addiction* 2005; **100**: 612–618.
 - 20 Fergusson DM, Horwood LJ, Ridder EM. Tests of causal linkages between cannabis use and psychotic symptoms. *Addiction* 2005; **100**: 354–366.
 - 21 Compton MT, Furman AC, Kaslow NJ. Lower negative symptom scores among cannabis-dependent patients with schizophrenia-spectrum disorders: preliminary evidence from an African American first-episode sample. *Schizophr Res* 2004; **71**: 61–64.
 - 22 Peralta V, Cuesta MJ. Influence of cannabis abuse on schizophrenic psychopathology. *Acta Psychiatr Scand* 1992; **85**: 127–130.
 - 23 Dixon L, Haas G, Weiden PJ, Sweeney J, Frances AJ. Drug abuse in schizophrenic patients: clinical correlates and reasons for use. *Am J Psychiatry* 1991; **148**: 224–230.
 - 24 Henquet C, Di Forti M, Morrison P, Kuepper R, Murray RM. Gene-environment interplay between cannabis and psychosis. *Schizophr Bull* 2008; **34**: 1111–1121.
 - 25 Sullivan PF, Kendler KS, Neale MC. Schizophrenia as a complex trait—Evidence from a meta-analysis of twin studies. *Arch Gen Psychiatry* 2003; **60**: 1187–1192.
 - 26 Ripke S, Sanders AR, Kendler KS, Levinson DF, Sklar P, Holmans PA et al. Genome-wide association study identifies five new schizophrenia loci. *Nat Genet* 2011; **43**: 969–976.
 - 27 Ripke S, O'Dushlaine C, Chambert K, Moran JL, Kahler AK, Akterin S et al. Genome-wide association analysis identifies 13 new risk loci for schizophrenia. *Nat Genet* 2013; **45**: 1150–1159.
 - 28 Kendler KS, Schmitt E, Aggen SH, Prescott CA. Genetic and environmental influences on alcohol, caffeine, cannabis, and nicotine use from early adolescence to middle adulthood. *Arch Gen Psychiatry* 2008; **65**: 674–682.
 - 29 Agrawal A, Lynskey MT. The genetic epidemiology of cannabis use, abuse and dependence. *Addiction* 2006; **101**: 801–812.
 - 30 Verweij KJ, Zietsch BP, Lynskey MT, Medland SE, Neale MC, Martin NG et al. Genetic and environmental influences on cannabis use initiation and problematic use: a meta-analysis of twin studies. *Addiction* 2010; **105**: 417–430.
 - 31 Verweij KJ, Vinkhuyzen AA, Benyamin B, Lynskey MT, Quaye L, Agrawal A et al. The genetic aetiology of cannabis use initiation: a meta-analysis of genome-wide association studies and a SNP-based heritability estimation. *Addict Biol* 2013; **18**: 846–850.
 - 32 Purcell SM, Wray NR, Stone JL, Visscher PM, O'Donovan MC, Sullivan PF et al. Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. *Nature* 2009; **460**: 748–752.
 - 33 Bucholz KK, Heath AC, Madden PA. Transitions in drinking in adolescent females: evidence from the Missouri adolescent female twin study. *Alcohol Clin Exp Res* 2000; **24**: 914–923.
 - 34 Knopik VS, Heath AC, Madden PA, Bucholz KK, Slutske WS, Nelson EC et al. Genetic effects on alcohol dependence risk: re-evaluating the importance of psychiatric and other heritable risk factors. *Psychol Med* 2004; **34**: 1519–1530.
 - 35 Heath AC, Whitfield JB, Martin NG, Pergadia ML, Goate AM, Lind PA et al. A quantitative-trait genome-wide association study of alcoholism risk in the community: findings and implications. *Biol Psychiatry* 2011; **70**: 513–518.
 - 36 Medland SE, Nyholt DR, Painter JN, McEvoy BP, McRae AF, Zhu G et al. Common variants in the trichohyalin gene are associated with straight hair in Europeans. *Am J Hum Genet* 2009; **85**: 750–755.
 - 37 Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D et al. PLINK: A tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 2007; **81**: 559–575.
 - 38 StataCorp. Stata Statistical Software: Release 12. StataCorp LP: College Station, TX, USA, 2011.
 - 39 Smoller JW, Craddock N, Kendler K, Lee PH, Neale BM, Nurnberger JI et al. Identification of risk loci with shared effects on five major psychiatric disorders: a genome-wide analysis. *Lancet* 2013; **381**: 1371–1379.
 - 40 Foti DJ, Kotov R, Guey LT, Bromet EJ. Cannabis use and the course of schizophrenia: 10-year follow-up after first hospitalization. *Am J Psychiatry* 2010; **167**: 987–993.
 - 41 Power RA, Wingenbach T, Cohen-Woods S, Uher R, Ng MY, Butler AW et al. Estimating the heritability of reporting stressful life events captured by common genetic variants. *Psychol Med* 2012; **43**: 1965–1971.
 - 42 Whitfield JB, Zhu G, Heath AC, Martin NG. Choice of residential location: chance, family influences, or genes? *Twin Res Hum Genet* 2005; **8**: 22–26.
 - 43 Winter T, Kaprio J, Viken RJ, Karvonen S, Rose RJ. Individual differences in adolescent religiosity in Finland: familial effects are modified by sex and region of residence. *Twin Res* 1999; **2**: 108–114.
 - 44 Hickman M, Vickerman P, Macleod J, Lewis G, Zammit S, Kirkbride J et al. If cannabis caused schizophrenia—how many cannabis users may need to be prevented in order to prevent one case of schizophrenia? England and Wales calculations. *Addiction* 2009; **104**: 1856–1861.

Medical Marijuana for Digestive Disorders: High Time to Prescribe?

Mark E. Gerich, MD^{1,4}, Robert W. Isfort, MD^{1,4}, Bryan Brimhall, MD² and Corey A. Siegel, MD³

The use of recreational and medical marijuana is increasingly accepted by the general public in the United States. Along with growing interest in marijuana use has come an understanding of marijuana's effects on normal physiology and disease, primarily through elucidation of the human endocannabinoid system. Scientific inquiry into this system has indicated potential roles for marijuana in the modulation of gastrointestinal symptoms and disease. Some patients with gastrointestinal disorders already turn to marijuana for symptomatic relief, often without a clear understanding of the risks and benefits of marijuana for their condition. Unfortunately, that lack of understanding is shared by health-care providers. Marijuana's federal legal status as a Schedule I controlled substance has limited clinical investigation of its effects. There are also potential legal ramifications for physicians who provide recommendations for marijuana for their patients. Despite these constraints, as an increasing number of patients consider marijuana as a potential therapy for their digestive disorders, health-care providers will be asked to discuss the issues surrounding medical marijuana with their patients.

SUPPLEMENTARY MATERIAL is linked to the online version of the paper at <http://www.nature.com/ajg>

Am J Gastroenterol 2015; 110:208–214; doi:10.1038/ajg.2014.245; published online 9 September 2014

Archaeological records indicate that marijuana, also known as cannabis, has been cultivated and used for its psychoactive and medicinal properties for over 2700 years (1). With the Controlled Substances Act of 1971, however, marijuana was classified in the United States as a Schedule I substance with no accepted medical use and high potential for abuse, similar to heroin (2). Despite federal prohibition, the use of marijuana for recreational and medicinal purposes continues to increase (3,4). Following the 2013 election cycle, medical marijuana programs exist in 21 states as well as Washington DC. Furthermore, the recreational use of marijuana was legalized in both Colorado and the state of Washington in 2012 and it has been decriminalized in 15 additional states (**Figure 1**). With these developments, medical professionals who care for patients with digestive disorders are increasingly faced with questions about the therapeutic role of marijuana and, in some cases, are asked to provide documentation to support a request for medical marijuana. This brief review is intended to help inform those discussions.

Phytocannabinoids and the endocannabinoid system

Marijuana is the common name for the *Cannabis* plant, from which nearly 500 different chemical compounds have been

isolated (5). Among these, the most clinically relevant are the phytocannabinoids that are concentrated in the plant's flowering buds that are harvested for consumption. The vast majority of interest has focused on Δ^9 -tetrahydrocannabinol (THC), which is primarily responsible for the psychoactive effects of marijuana (6). Marijuana also contains ~70 other phytocannabinoids, such as cannabidiol (CBD), that are present in varying ratios when compared with THC content and seem to have minimal psychotropic effects (5,7). There are two main subspecies of the *Cannabis* plant: *Cannabis sativa* and *Cannabis indica*. *Sativa*-dominant strains have higher THC content than *indica* strains, in which the CBD content is higher (8).

Scientific interest in the medical application of marijuana-based compounds heightened in the early 1990s with the discovery of an endogenous cannabinoid signaling system, termed the endocannabinoid system, through which phytocannabinoids appear to signal. The endocannabinoid system has since been implicated in diverse physiologic processes (9). It includes two G protein-coupled cannabinoid receptors, CB1 and CB2, as well as two endogenous ligands or endocannabinoids: anandamide and 2-arachidonylglycerol (10). Generally, CB1 is found in greatest abundance in central and peripheral neurons, whereas CB2

¹Division of Gastroenterology and Hepatology, University of Colorado, Aurora, Colorado, USA; ²Department of Medicine, University of Colorado, Aurora, Colorado, USA; ³Inflammatory Bowel Disease Center, Dartmouth-Hitchcock Medical Center, Lebanon, New Hampshire, USA; ⁴The first two authors contributed equally to this work. **Correspondence:** Mark E. Gerich, MD, Division of Gastroenterology and Hepatology, University of Colorado Anschutz Medical Campus, 12700 E. 19th Avenue, MS B-146, Aurora, Colorado 80045, USA. E-mail: mark.gerich@ucdenver.edu

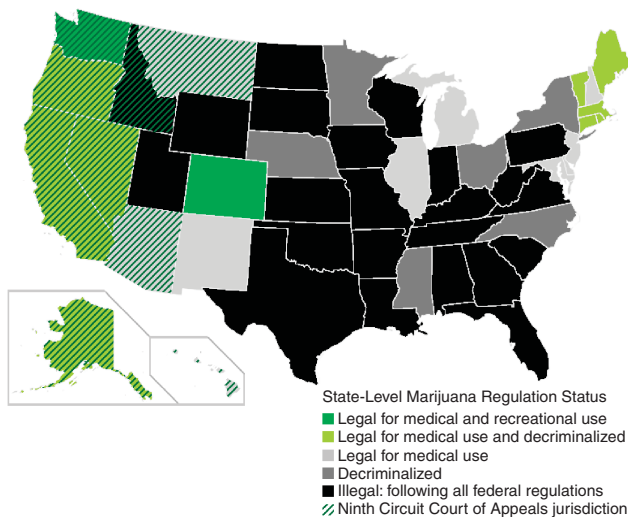


Figure 1. State-level marijuana regulation status.

is expressed predominantly in immune cells (11,12). In the gastrointestinal tract, CB1 receptors are expressed principally in the enteric nervous system with high concentration within cholinergic neurons of the myenteric and submucosal plexus where receptors are thought to promote an inhibitory effect on motility and secretory function via reduced acetylcholine release (13). CB1 receptors have also been identified in normal human colonic epithelium and smooth muscle, and CB1 receptor activation has been shown to enhance epithelial wound healing. CB2 receptor expression appears to be more pronounced in inflamed colonic epithelium and lamina propria immune cells, and there is *in vitro* evidence to suggest that activation of epithelial CB2 receptors by cannabinoids inhibits tumor necrosis factor- α -induced interleukin-8 release (14,15). Expression of cannabinoid receptors is very limited in the normal liver but is increased in experimental liver injury and cirrhosis. CB1 and CB2 receptor activation have been shown to induce pro- and anti-fibrogenic effects, respectively. In the normal pancreas, there is weak expression of CB1 and CB2 that increases in the setting of inflammation; however, there have been conflicting studies on the impact of cannabinoid receptor activation on experimental pancreatitis. There is also evidence that cannabinoid agonists have apoptotic, antiproliferative, and antimetastatic effects in several gastrointestinal cancer cell lines and animal models (16).

THC is a partial agonist of both CB1 and CB2 but has higher affinity for CB1, which appears to mediate its psychoactive properties. CBD has much weaker affinity for cannabinoid receptors but has demonstrated anti-inflammatory properties that may occur through CB2 inverse agonism or independently of cannabinoid receptors altogether (17). Several synthetic cannabinoid compounds have been developed to try to modulate the endocannabinoid system for therapeutic purposes. Two have been approved by the Food and Drug Administration (FDA). Dronabinol (Marinol) is a synthetic THC that is indicated for chemotherapy-induced nausea and vomiting as well as AIDS anorexia. Another THC

analog, known as nabilone (Cesamet), was approved for nausea and vomiting after cancer chemotherapy unresponsive to typical antiemetics. A third commercial medication, nabiximols (Sativex), is an oromucosal spray with relatively equal amounts of THC and CBD. It is approved in 13 countries, including Canada and the United Kingdom, for use in patients with cancer pain, neuropathic pain, or spasticity due to multiple sclerosis. Rimonabant (Acomplia) is a selective CB1 antagonist that was marketed for weight loss and used for smoking cessation but has been withdrawn because of adverse psychiatric effects—primarily depression. Numerous other CB1 and CB2 agonists and antagonists are currently undergoing investigation, including a phase II study in Europe of an oral therapy for ulcerative colitis that contains CBD and THC in a ratio of 20:1 (18,19).

Practical aspects and consequences of medical marijuana use

The psychotropic and physiologic effects of marijuana can vary greatly for different individuals depending on the route of administration, the relative dosage of THC and other phytocannabinoids, and the chronicity of use (20). An array of medical marijuana products is now available, including many different edible forms—from brownies and honey to barbecue sauce and soda—as well as very potent concentrates. Still, the most common method of consumption remains smoking that is done through a variety of delivery devices. These include marijuana cigarettes (“joints”), pipes (“bowls”), or water pipes with a chamber for water filtration (“bongs”). Vaporizers are often used to heat marijuana to a temperature sufficient to evaporate cannabinoids but not burn plant material, resulting in limited inhalation of tar and irritants (21). Other new inhalation methods, such as water pipes with carbon filtration systems, appear to reduce exposure to pesticides present in cultivated marijuana (22).

Product labeling and testing requirements are not standardized across states but they commonly include information about pesticides and contaminants as well as THC potency. State-licensed testing facilities often also provide data on the percentage of other non-THC phytocannabinoids such as CBD, cannabinol, cannabigerol, cannabichromene, and tetrahydrocannabivarin. Despite a lack of clinical data, statements regarding the efficacy for these compounds are often communicated to would-be buyers (23,24). The relative cannabinoid composition and THC content of marijuana products varies greatly, with THC content typically in the range of 0.5 to 5%; however, it has been increasing over time and concentrates can contain over 50% THC (25,26).

Following inhalation, the psychotropic effects of marijuana generally start within a minute, peak within a half hour, and begin tapering within 2–3 h. Following oral ingestion, physiological effects typically begin after 30–90 min, peak after 2–3 h, and can last 4–12 h, depending on dose and specific effect (27). Notably, this delayed onset can lead to difficulty in regulating dosage to achieve a desired effect in real time.

Acute adverse effects include anxiety and psychotic symptoms. The most concerning adverse effects of chronic use include

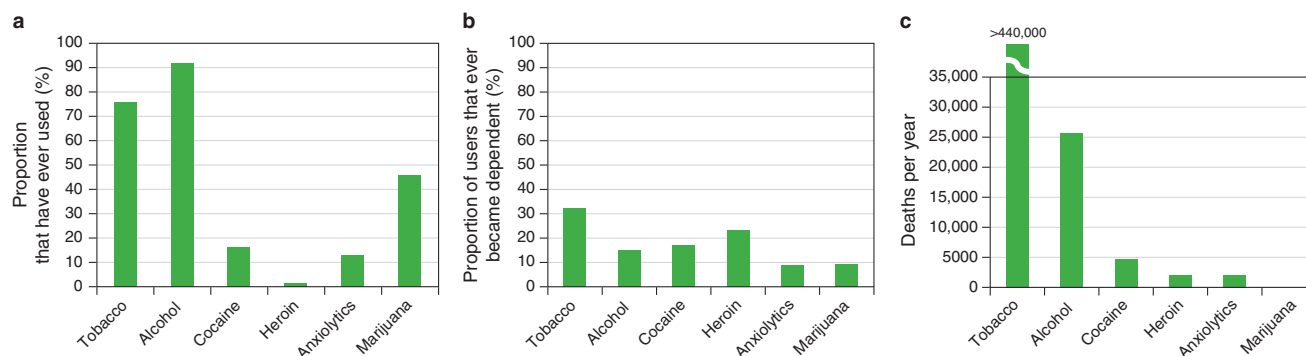


Figure 2. Drug use, dependency, and mortality statistics for the United States. The proportion of US individuals aged 15 to 54 years who (a) ever used common drugs (31), and (b) the proportion of drug users who ever became dependent (31). (c) Annual deaths in the United States attributable to tobacco (32), alcohol (33), cocaine (34), heroin (34), anxiolytics (34), and marijuana (35).

increased risk of motor vehicle crashes, decreased fertility, altered adolescent psychosocial development and mental health, as well as a hyperemesis syndrome characterized by cyclic episodes of nausea and vomiting along with a learned behavior of frequent hot bathing to relieve symptoms (28–31). Smoked crude marijuana contains similar carcinogens as tobacco smoke, with up to three times the tar content; however, long-term daily use of small amounts of marijuana does not appear to affect pulmonary function, and marijuana use has been associated with neither lung cancer nor head and neck cancer (32–34). With that said, it should be noted that the illegal federal status of marijuana has enabled far less rigorous safety evaluation for marijuana than for tobacco. There is mounting evidence of marijuana dependence as well as a withdrawal syndrome that is similar to tobacco withdrawal and includes irritability, sleep disturbance, anorexia, and depressed mood (35). A fatal dose of THC is *estimated* at 15–70 g based on studies in rodents; however, this is much higher than can be reached by heavy marijuana users and no deaths have been solely attributed to the use of marijuana (Figure 2) (28,36–41). Given the many variables involved, the actual dose of THC delivered to a user is not easily quantified and some experts suggest that patients should optimize their own dose of marijuana to achieve a desired effect rather than follow a prescription (42).

Efficacy for digestive disorders

Numerous preclinical studies indicate that endocannabinoids are involved in many functions in the digestive system, including gastric acid production, nausea and emesis, food intake, visceral sensation, gastrointestinal motility, hepatic fibrogenesis, and intestinal inflammation (10). Although modulation of the endocannabinoid system has shown therapeutic potential in a variety of experimental models of gastrointestinal disease (16,43–48), a vast majority of data from controlled human studies relate to synthetic cannabinoids. There is very limited clinical evidence demonstrating either a beneficial or detrimental effect of medical marijuana for digestive disorders. Through a systematic search of the existing scientific literature, we identified only five randomized controlled trials evaluating the impact of

marijuana on gastrointestinal function, symptoms, or disease (Table 1) (49–53).

Nausea and vomiting. The oral THC analogs (dronabinol and nabilone) have compared favorably with dopamine antagonists for chemotherapy-induced nausea and vomiting, although there are little data comparing them with 5-HT₃ antagonists or the NK₁ antagonist, aprepitant (20,54–56). Unlike chemotherapy-induced vomiting, the associated nausea has been less responsive to current first-line therapies, and considerable preclinical evidence indicates that manipulation of the endocannabinoid system may be beneficial in some cases (57). Although many patients have a strong preference for smoked marijuana as a therapy for nausea, there are no controlled studies evaluating the antiemetic effects of marijuana (54).

Appetite stimulation and weight gain. Marijuana is known as an appetite stimulant. This impression has been borne out by studies demonstrating that healthy subjects who smoke marijuana have higher food consumption, caloric intake, and body weight (49,58). Nevertheless, among cancer patients with anorexia and/or cachexia, there are mixed data from controlled trials to suggest a modest benefit, at best, from synthetic THC (dronabinol) (59–65). For patients with AIDS-related anorexia, there is evidence to suggest that dronabinol improves anorexia and, at higher doses, leads to weight gain, although results have varied (50,52,66–68). Two small, placebo-controlled studies in this population demonstrated that smoking marijuana (2–4% THC) three times per week increased food intake and body weight, with greater increases if marijuana was smoked four times daily (50,52). There are currently no data from controlled studies to indicate a significant benefit from ingested marijuana extract (69) or smoked marijuana for cancer-related anorexia and/or cachexia.

Hepatitis C virus infection. Daily marijuana use among patients with hepatitis C virus infection has been associated with increased steatosis and fibrosis (70–72). Conversely, two relatively small studies evaluated the impact of synthetic

Table 1. Randomized controlled trials evaluating the impact of marijuana on gastrointestinal function, symptoms, or disease

First author	Year	Function or disorder studied	Treatment	Subjects	Outcome	Reference
Foltin	1988	Appetite	Smoked marijuana (2.3% THC); placebo	6 Healthy adult males	40% Increase in daily caloric intake because of more frequent snacking with marijuana	(49)
Haney	2005	HIV	Smoked marijuana (1.8, 2.9, 3.9% THC); dronabinol; placebo	30 HIV+ patients	Comparable increases in caloric intake for marijuana and dronabinol over placebo	(50)
Strasser	2006	Cancer-related anorexia-cachexia	Oral cannabis extract (2.5 mg THC, 1 mg CBD); oral THC; placebo	164 Cancer patients	No difference in appetite or quality-of-life outcomes	(51)
Haney	2007	HIV	Smoked marijuana (2.0, 3.9% THC); dronabinol; placebo	10 HIV+ patients	Dose-dependent increase in caloric intake and body weight for marijuana and dronabinol over placebo	(52)
Naftali	2013	Crohn's disease	<i>Cannabis sativa</i> cigarette (23% THC, 0.5% CBD); placebo	21 Patients with moderately active Crohn's disease	Significant clinical response but no decrease in inflammatory markers with marijuana	(53)

CBD, cannabidiol; THC, tetrahydrocannabinol.
See **Appendix** for systematic review methodology.

cannabinoid or marijuana use during interferon-based therapy and suggested that they may increase adherence and sustained virologic response, presumably through a reduction in treatment-associated symptoms (73,74).

Inflammatory bowel disease (IBD). Patients with both ulcerative colitis and Crohn's disease commonly use marijuana for relief of symptoms (75). The majority of IBD patients report that marijuana provides significant benefit for poor appetite, nausea, and abdominal pain; however, improvement in diarrhea is less clear (76). Notably, if it were available legally, over half of nonusers have reported an interest in using marijuana for IBD symptom relief (76). Small retrospective and prospective observational studies of patients with Crohn's disease who smoke medical marijuana that contains moderate concentrations of THC indicate that medical marijuana has beneficial effects on symptom-based disease activity indices, overall well-being, and steroid use (77,78). A small randomized controlled trial of marijuana smoked twice daily (115 mg THC; 23% THC and <0.5% CBD) also recently demonstrated a significant response as measured by the Crohn's Disease Activity Index (53). Endoscopic disease activity was not evaluated, but lack of reduction in C-reactive protein indicates that the inflammatory disease burden was not affected. Although patients with IBD may feel better while using marijuana, evidence for an objective improvement in disease activity is lacking.

Abdominal pain. Preclinical evidence indicates that the endocannabinoid system plays an important role in the modulation of pain. Although there is also significant evidence that synthetic cannabinoids and medical marijuana modulate chronic pain, especially chronic neuropathic pain, there are no controlled

human studies specifically evaluating the efficacy of medical marijuana for chronic abdominal pain (20).

Marijuana, physicians, and the law

State medical marijuana laws vary widely with respect to the medical conditions for which marijuana is approved, the administrative steps necessary for patients to legally obtain marijuana, and the role physicians must play in the process (**Supplementary Table S1** online). Cancer and HIV/AIDS are considered qualifying medical conditions in all states with defined medical marijuana programs. Many states also permit the use of marijuana for other conditions including nausea, cachexia, hepatitis C infection, and Crohn's disease.

Most states that allow for medical marijuana have designed systems that seem to protect physicians. Given the federal prohibition against writing a prescription for medical marijuana, state laws commonly direct physicians to provide patients with documents termed "certifications," "recommendations," or "referrals" for marijuana use. Generally, before providing certification of medical necessity for marijuana to a patient, a physician must have an active state medical license and have an ongoing therapeutic relationship with that patient, the definition of which varies by state. Some states also require specific certification to recommend medical marijuana (see **Supplementary Table S1**). Physicians are not allowed to have contact with marijuana dispensaries, to refer patients to a specific dispensary, to have a financial relationship with a dispensary, or to certify family members or themselves for marijuana use.

Still, in addition to concerns about safety and efficacy, the federal legal status of marijuana as a Schedule I controlled substance leads many physicians to be wary of recommending

medical marijuana to their patients. Currently, the possession, manufacture, and distribution of marijuana—even for purposes of medical treatment—remain a violation of the federal Controlled Substance Act. It is not entirely clear whether a physician would be liable under federal law for the act of recommending or prescribing marijuana for a patient's medical condition. In response to state initiatives that attempted to shield physicians who recommended the medicinal use of marijuana, the federal government declared a policy that "recommending or prescribing Schedule I controlled substances" would lead to revocation of a physician's registration to prescribe controlled substances, exclusion from the Medicare/Medicaid programs, and criminal prosecution (79). A group of physicians and patients subsequently obtained a permanent injunction against enforcement of this policy (80). In the case of *Conant vs. Walters*, the Ninth Circuit Court of Appeals then upheld the injunction on the grounds that the federal government's drug enforcement policy violates the First Amendment values of the physician–patient relationship (81). When the federal government then appealed the *Conant* decision to the US Supreme Court, it declined to hear the case at that time; it may still hear future appeals.

For physicians in states within the Ninth Circuit's jurisdiction (see **Figure 1**), this means that the federal government continues to be forbidden from prosecuting physicians solely on the basis of their communication with patients regarding marijuana. However, for physicians in other jurisdictions, *Conant* is not a binding precedent and the threat of federal prosecution for recommending medical marijuana remains. Under the Obama administration, the Department of Justice has indicated several times that federal resources will generally not be used to prosecute individuals or caregivers who are in compliance with state marijuana laws (82,83). Nevertheless, the Department of Justice retains the authority to enforce federal marijuana laws. Although seemingly unlikely, it remains to be seen whether a new federal administration will choose to enforce federal law and prosecute physicians who recommend medical marijuana, notwithstanding a state's law to the contrary.

Concluding remarks

It is increasingly clear that the endocannabinoid system plays a role in diverse biological pathways that affect gastrointestinal and hepatic physiology and pathology. However, due in large part to a dearth of high-quality human study data, the clinical efficacy of marijuana or its constituent phytocannabinoids for digestive disorders remains unclear. Although there are no studies proving the long-term safety of marijuana use, its safety profile compares favorably to other illicit substances, to legal intoxicants like alcohol, possibly to opioids, and to some existing therapies for digestive disorders. Claims of marijuana as a cure-all are clearly unfounded but, at the least, it appears to hold promise as a modifier of gastrointestinal symptoms. As medical marijuana use continues to grow in the United States, physicians must take the lead in understanding the risks and benefits in order to provide accurate information to patients. This understanding necessitates

further well-designed scientific inquiry that will require federal reclassification of marijuana from its current status as a Schedule I controlled substance. In the meantime, physicians are in a challenging position between providing patients with their opinion and minimizing punitive consequences.

ACKNOWLEDGMENTS

Chris Puckett (Associate Counsel, University of Colorado Hospital) provided voluntary input on legal issues surrounding marijuana recommendations.

CONFLICT OF INTEREST

Guarantor of the article: Mark E. Gerich, MD.

Specific author contributions: Mark E. Gerich: paper concept, literature review, writing, and proofreading/revision; Robert W. Isfort: literature review, writing, proofreading, and table design; Bryan Brimhall: literature review, graphics design, and writing; Corey A. Siegel: paper concept and proofreading/revision.

Financial support: None.

Potential competing interests: None.

Study Highlights

WHAT IS CURRENT KNOWLEDGE

- ✓ Despite federal prohibition, the use of marijuana for recreational and medicinal purposes continues to increase.
- ✓ Some patients with gastrointestinal disorders self-medicate with marijuana.
- ✓ Increasing marijuana acceptance is generating more questions regarding its use for health conditions.

WHAT IS NEW HERE

- ✓ A focused summary of the data to support the use of marijuana for digestive disorders.
- ✓ An update on the legal considerations for healthcare providers who are considering recommending medical marijuana for patients.

REFERENCES

1. Russo EB, Jiang HE, Li X *et al*. Phytochemical and genetic analyses of ancient cannabis from Central Asia. *J Exp Bot* 2008;59:4171–82.
2. Mack A, Joy JE. Marijuana as Medicine?: The Science Beyond the Controversy National Academy Press: Washington, DC, 2001.
3. Degenhardt L, Chiu WT, Sampson N *et al*. Toward a global view of alcohol, tobacco, cannabis, and cocaine use: findings from the WHO World Mental Health Surveys. *PLoS Med* 2008;5:e141.
4. Bowles DW. Persons registered for medical marijuana in the United States. *J Palliat Med* 2012;15:9–11.
5. Elsohly MA, Slade D. Chemical constituents of marijuana: the complex mixture of natural cannabinoids. *Life Sci* 2005;78:539–48.
6. Gaoni Y, Mechoulam R. Isolation, structure, and partial synthesis of an active constituent of hashish. *J Am Chem Soc* 1964;86:1646–7.
7. Ashton CH. Pharmacology and effects of cannabis: a brief review. *Br J Psychiatry* 2001;178:101–6.
8. Atakan Z. Cannabis, a complex plant: different compounds and different effects on individuals. *Ther Adv Psychopharmacol* 2012;2:241–54.
9. Klein TW. Cannabinoid-based drugs as anti-inflammatory therapeutics. *Nat Rev Immunol* 2005;5:400–11.
10. Izzo AA, Sharkey KA. Cannabinoids and the gut: new developments and emerging concepts. *Pharmacol Ther* 2010;126:21–38.

11. Matsuda LA, Lolait SJ, Brownstein MJ *et al*. Structure of a cannabinoid receptor and functional expression of the cloned cDNA. *Nature* 1990;346:561–4.
12. Munro S, Thomas KL, Abu-Shaar M. Molecular characterization of a peripheral receptor for cannabinoids. *Nature* 1993;365:61–5.
13. Di Carlo G, Izzo AA. Cannabinoids for gastrointestinal diseases: potential therapeutic applications. *Expert Opin Investig Drugs* 2003;12:39–49.
14. Wright K, Rooney N, Feeney M *et al*. Differential expression of cannabinoid receptors in the human colon: cannabinoids promote epithelial wound healing. *Gastroenterology* 2005;129:437–53.
15. Ihenetu K, Molleman A, Parsons ME *et al*. Inhibition of interleukin-8 release in the human colonic epithelial cell line HT-29 by cannabinoids. *Eur J Pharmacol* 2003;458:207–15.
16. Izzo AA, Camilleri M. Emerging role of cannabinoids in gastrointestinal and liver diseases: basic and clinical aspects. *Gut* 2008;57:1140–55.
17. Pertwee RG. The diverse CB1 and CB2 receptor pharmacology of three plant cannabinoids: delta9-tetrahydrocannabinol, cannabidiol and delta9-tetrahydrocannabivarin. *Br J Pharmacol* 2008;153:199–215.
18. Greinisen WE, Turner H. Immunoreactive effects of cannabinoids: considerations for the therapeutic use of cannabinoid receptor agonists and antagonists. *Int Immunopharmacol* 2010;10:547–55.
19. A pilot study of GWP42003 in the symptomatic treatment of ulcerative colitis (GWID10160); <http://clinicaltrials.gov/show/NCT01562314>; Accessed 7 March 2014.
20. Information for Healthcare Professionals: Cannabis (marihuana, marijuana) and the cannabinoids. Health Canada <http://www.hc-sc.gc.ca/dhp-mps/marihuana/med/infoprof-eng.php>; Accessed 13 February 2014.
21. Hazekamp A, Ruhaak R, Zuurman L *et al*. Evaluation of a vaporizing device (Volcano) for the pulmonary administration of tetrahydrocannabinol. *J Pharm Sci* 2006;95:1308–17.
22. Sullivan N, Elzinga S, Raber JC. Determination of pesticide residues in cannabis smoke. *J Toxicol* 2013;2013:378168.
23. Izzo AA, Borrelli F, Capasso R *et al*. Non-psychoactive plant cannabinoids: new therapeutic opportunities from an ancient herb. *Trends Pharmacol Sci* 2009;30:515–27.
24. Russo EB. Taming THC: potential cannabis synergy and phytocannabinoid-terpenoid entourage effects. *Br J Pharmacol* 2011;163:1344–64.
25. Cannabis: A health perspective and research agenda. World Health Organization Division of Mental Health and Prevention of Substance Abuse. WHO/MSA/PSA/97.4. World Health Organization: 1997.
26. McLaren J, Swift W, Dillon P *et al*. Cannabis potency and contamination: a review of the literature. *Addiction* 2008;103:1100–9.
27. Grotenhermen F. Pharmacokinetics and pharmacodynamics of cannabinoids. *Clin Pharmacokinet* 2003;42:327–60.
28. Hall W, Degenhardt L. Adverse health effects of non-medical cannabis use. *Lancet* 2009;374:1383–91.
29. Allen JH, de Moore GM, Heddle R *et al*. Cannabinoid hyperemesis: cyclical hyperemesis in association with chronic cannabis abuse. *Gut* 2004;53:1566–70.
30. Simonetto DA, Oxentenko AS, Herman ML *et al*. Cannabinoid hyperemesis: a case series of 98 patients. *Mayo Clin Proc* 2012;87:114–9.
31. Nicolson SE, Denysenko L, Mulcare JL *et al*. Cannabinoid hyperemesis syndrome: a case series and review of previous reports. *Psychosomatics* 2012;53:212–9.
32. Bowles DW, O'Bryant CL, Camidge DR *et al*. The intersection between cannabis and cancer in the United States. *Crit Rev Oncol Hematol* 2012;83:1–10.
33. Hashibe M, Morgenstern H, Cui Y *et al*. Marijuana use and the risk of lung and upper aerodigestive tract cancers: results of a population-based case-control study. *Cancer Epidemiol Biomarkers Prev* 2006;15:1829–34.
34. Pletcher MJ, Vittinghoff E, Kalhan R *et al*. Association between marijuana exposure and pulmonary function over 20 years. *JAMA* 2012;307:173–81.
35. Hesse M, Thylstrup B. Time-course of the DSM-5 cannabis withdrawal symptoms in poly-substance abusers. *BMC Psychiatry* 2013;13:258.
36. Administration DE. Drugs of abuse. A DEA Resource Guide: DEA, 2011: 68–69.
37. Warner LA, Kessler RC, Hughes M *et al*. Prevalence and correlates of drug use and dependence in the United States. Results from the National Comorbidity Survey. *Arch Gen Psychiatry* 1995;52:219–29.
38. Smoking-attributable mortality, years of potential life lost, and productivity losses--United States, 2000–2004. *MMWR Morb Mortal Wkly Rep* 2008;57:1226–8.
39. Murphy SL, Xu J, Kochanek KD. Deaths: final data for 2010. *Natl Vital Stat Rep* 2013;61:1–118.
40. Minino AM, Anderson RN, Fingerhut LA *et al*. Deaths: injuries, 2002. *Natl Vital Stat Rep* 2006;54:1–124.
41. Gable RS. The toxicity of recreational drugs. *Am Scientist* 2006;94:206–8.
42. Aggarwal SK, Kyashna-Tocha M, Carter GT. Dosing medical marijuana: rational guidelines on trial in Washington State. *Med Gen Med* 2007;9:52.
43. Wright KL, Duncan M, Sharkey KA. Cannabinoid CB2 receptors in the gastrointestinal tract: a regulatory system in states of inflammation. *Br J Pharmacol* 2008;153:263–70.
44. Julien B, Grenard P, Teixeira-Clerc F *et al*. Antifibrogenic role of the cannabinoid receptor CB2 in the liver. *Gastroenterology* 2005;128:742–55.
45. Teixeira-Clerc F, Julien B, Grenard P *et al*. CB1 cannabinoid receptor antagonism: a new strategy for the treatment of liver fibrosis. *Nat Med* 2006;12:671–6.
46. Storr MA, Keenan CM, Zhang H *et al*. Activation of the cannabinoid 2 receptor (CB2) protects against experimental colitis. *Inflamm Bowel Dis* 2009;15:1678–85.
47. Mallat A, Teixeira-Clerc F, Deveaux V *et al*. The endocannabinoid system as a key mediator during liver diseases: new insights and therapeutic openings. *Br J Pharmacol* 2011;163:1432–40.
48. Mallat A, Teixeira-Clerc F, Lotersztajn S. Cannabinoid signaling and liver therapeutics. *J Hepatol* 2013;59:891–6.
49. Foltin RW, Fischman MW, Byrne MF. Effects of smoked marijuana on food intake and body weight of humans living in a residential laboratory. *Appetite* 1988;11:1–14.
50. Haney M, Rabkin J, Gunderson E *et al*. Dronabinol and marijuana in HIV(+) marijuana smokers: acute effects on caloric intake and mood. *Psychopharmacology (Berl)* 2005;181:170–8.
51. Strasser F, Luftner D, Possinger K *et al*. Comparison of orally administered cannabis extract and delta-9-tetrahydrocannabinol in treating patients with cancer-related anorexia-cachexia syndrome: a multicenter, phase III, randomized, double-blind, placebo-controlled clinical trial from the Cannabis-In-Cachexia-Study-Group. *J Clin Oncol* 2006;24:3394–400.
52. Haney M, Gunderson EW, Rabkin J *et al*. Dronabinol and marijuana in HIV-positive marijuana smokers. Caloric intake, mood, and sleep. *J Acquir Immune Defic Syndr* 2007;45:545–54.
53. Naftali T, Bar-Lev Schleider L, Dotan I *et al*. Cannabis induces a clinical response in patients with Crohn's disease: a prospective placebo-controlled study. *Clin Gastroenterol Hepatol* 2013;11:1276–80.e1.
54. Tramer MR, Carroll D, Campbell FA *et al*. Cannabinoids for control of chemotherapy induced nausea and vomiting: quantitative systematic review. *BMJ* 2001;323:16–21.
55. Machado Rocha FC, Stefano SC, De Cassia Haiek R *et al*. Therapeutic use of Cannabis sativa on chemotherapy-induced nausea and vomiting among cancer patients: systematic review and meta-analysis. *Eur J Cancer Care (Engl)* 2008;17:431–43.
56. Meiri E, Jhangiani H, Vredenburg JJ *et al*. Efficacy of dronabinol alone and in combination with ondansetron versus ondansetron alone for delayed chemotherapy-induced nausea and vomiting. *Curr Med Res Opin* 2007;23:533–43.
57. Parker LA, Rock EM, Limebeer CL. Regulation of nausea and vomiting by cannabinoids. *Br J Pharmacol* 2011;163:1411–22.
58. Mattes RD, Engelman K, Shaw LM *et al*. Cannabinoids and appetite stimulation. *Pharmacol Biochem Behav* 1994;49:187–95.
59. Ekert H, Waters KD, Jurk IH *et al*. Amelioration of cancer chemotherapy-induced nausea and vomiting by delta-9-tetrahydrocannabinol. *Med J Aust* 1979;2:657–9.
60. Sallan SE, Cronin C, Zelen M *et al*. Antiemetics in patients receiving chemotherapy for cancer: a randomized comparison of delta-9-tetrahydrocannabinol and prochlorperazine. *N Engl J Med* 1980;302:135–8.
61. Plasse TF, Gorter RW, Krasnow SH *et al*. Recent clinical experience with dronabinol. *Pharmacol Biochem Behav* 1991;40:695–700.
62. Nelson K, Walsh D, Deeter P *et al*. A phase II study of delta-9-tetrahydrocannabinol for appetite stimulation in cancer-associated anorexia. *J Palliat Care* 1994;10:14–18.
63. Jatoi A, Windschitl HE, Loprinzi CL *et al*. Dronabinol versus megestrol acetate versus combination therapy for cancer-associated anorexia: a North Central Cancer Treatment Group study. *J Clin Oncol* 2002;20:567–73.
64. Mantovani G, Maccio A, Madeddu C *et al*. Randomized phase III clinical trial of five different arms of treatment in 332 patients with cancer cachexia. *Oncologist* 2010;15:200–11.

65. Brisbois TD, de Kock IH, Watanabe SM *et al*. Delta-9-tetrahydrocannabinol may palliate altered chemosensory perception in cancer patients: results of a randomized, double-blind, placebo-controlled pilot trial. *Ann Oncol* 2011;22:2086–93.
66. Beal JE, Olson R, Laubenstein L *et al*. Dronabinol as a treatment for anorexia associated with weight loss in patients with AIDS. *J Pain Symptom Manage* 1995;10:89–97.
67. Beal JE, Olson R, Lefkowitz L *et al*. Long-term efficacy and safety of dronabinol for acquired immunodeficiency syndrome-associated anorexia. *J Pain Symptom Manage* 1997;14:7–14.
68. Lutge EE, Gray A, Siegfried N. The medical use of cannabis for reducing morbidity and mortality in patients with HIV/AIDS. *Cochrane Database Syst Rev* 2013;4:CD005175.
69. Cannabis-In-Cachexia-Study-Group. Strasser F, Luftner D *et al*. Comparison of orally administered cannabis extract and delta-9-tetrahydrocannabinol in treating patients with cancer-related anorexia-cachexia syndrome: a multicenter, phase III, randomized, double-blind, placebo-controlled clinical trial from the Cannabis-In-Cachexia-Study-Group. *J Clin Oncol* 2006;24:3394–400.
70. Hezode C, Zafrani ES, Roudot-Thoraval F *et al*. Daily cannabis use: a novel risk factor of steatosis severity in patients with chronic hepatitis C. *Gastroenterology* 2008;134:432–9.
71. Hezode C, Roudot-Thoraval F, Nguyen S *et al*. Daily cannabis smoking as a risk factor for progression of fibrosis in chronic hepatitis C. *Hepatology* 2005;42:63–71.
72. Ishida JH, Peters MG, Jin C *et al*. Influence of cannabis use on severity of hepatitis C disease. *Clin Gastroenterol Hepatol* 2008;6:69–75.
73. Costiniuk CT, Mills E, Cooper CL. Evaluation of oral cannabinoid-containing medications for the management of interferon and ribavirin-induced anorexia, nausea and weight loss in patients treated for chronic hepatitis C virus. *Can J Gastroenterol* 2008;22:376–80.
74. Sylvestre DL, Clements BJ, Malibu Y. Cannabis use improves retention and virological outcomes in patients treated for hepatitis C. *Eur J Gastroenterol Hepatol* 2006;18:1057–63.
75. Lal S, Prasad N, Ryan M *et al*. Cannabis use amongst patients with inflammatory bowel disease. *Eur J Gastroenterol Hepatol* 2011;23:891–6.
76. Ravikoff Allegretti J, Courtwright A, Lucci M *et al*. Marijuana use patterns among patients with inflammatory bowel disease. *Inflamm Bowel Dis* 2013;19:2809–14.
77. Naftali T, Lev LB, Yablecovitch D *et al*. Treatment of Crohn's disease with cannabis: an observational study. *Isr Med Assoc J* 2011;13:455–8.
78. Lahat A, Lang A, Ben-Horin S. Impact of cannabis treatment on the quality of life, weight and clinical disease activity in inflammatory bowel disease patients: a pilot prospective study. *Digestion* 2012;85:1–8.
79. The Administration's Response to California Proposition 215 and Arizona Proposition 200, Office of National Drug Control Policy (30 December 1996).
80. 309 F.3d 629, 633 (9th Cir. 2002).
81. 309 F.3d 629, 639 (9th Cir. 2002).
82. Ogden D. Investigations and Prosecutions in States Authorizing the Medical Use of Marijuana. In: US Department of Justice Office of the Deputy Attorney General, ed 2009;3.
83. Cole J. Guidance Regarding Marijuana Enforcement. In: US Department of Justice Office of the Deputy Attorney General, ed. Washington DC, 2013.

APPENDIX

A systematic literature review was undertaken using the MEDLINE/PubMed database to identify randomized controlled clinical trials involving marijuana and gastrointestinal conditions. This was undertaken utilizing the search phrase: “cannabis OR marijuana OR tetrahydrocannabinol OR cannabidiol OR THC OR cannabinoid OR dronabinol OR nabilone OR nabiximols AND _____.” Results were filtered for human studies and clinical trials and were hand-reviewed for relevance. The final search term was variable and included the following: esophagus, stomach, pancreas, gallbladder, biliary, liver, small intestine, colon, large intestine, rectum, anus, cancer, esophageal cancer (adenocarcinoma), stomach cancer (adenocarcinoma), carcinoid, liver cancer, hepatocellular carcinoma, hepatoma, pancreatic

cancer (adenocarcinoma), neuroendocrine tumor, cholangiocarcinoma, gallbladder cancer, colon cancer (adenocarcinoma), rectal cancer (adenocarcinoma), anal cancer, polyp, polyposis, esophagitis, gastritis, ulcer, peptic ulcer disease, gastroesophageal reflux disease, heartburn, gastroesophageal reflux disease, pancreatitis, enteritis, colitis, cholecystitis, cholangitis, proctitis, inflammatory bowel disease, Crohn's disease, ulcerative colitis, regional enteritis, ulcerative proctitis, eosinophilic esophagitis, celiac disease, microscopic colitis, hepatitis, cirrhosis, cholelithiasis, gastroparesis, motility, gastric emptying, irritable bowel syndrome, dysphagia, odynophagia, dyspepsia, pyrosis, abdominal pain, nausea, vomiting, early satiety, diarrhea, constipation, ileus, weight, cachexia, anorexia, bulimia, and nutrition.